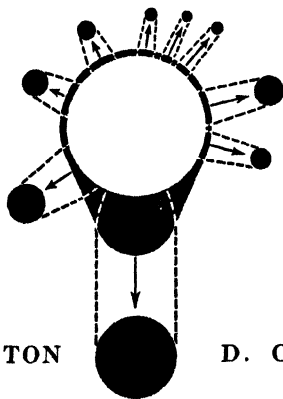# COLLOID SCIENCE

.

COLLOID

SCIENCE

by **JAMES W. McBAIN**

DIRECTOR, NATIONAL CHEMICAL LABORATORIES, POONA, INDIA

FORMERLY PROFESSOR OF CHEMISTRY, STANFORD UNIVERSITY

BOSTON D. C. HEATH AND COMPANY

DEDICATION

The author desires to place on record his most sincere appreciation of the perhaps unique action of the Bristol-Meyers Company in most generously endowing him with highly competent secretarial assistance, to enable him to produce this book and the numerous yearly publications of his collaborators.

FOREWORD

Today the prodigious and multifarious growth of Science and her almost innumerable ramifications make it hopeless for a single brain to master more than a tiny fraction of one of her minor subdivisions. A modern equivalent of so omniscient a scientist as Alexander von Humboldt, for example, is no longer a possibility; or, limiting oneself to chemistry, a contemporary knowledge of it as complete and encyclopedic as was that of Berzelius is now out of the question.

Nevertheless, there have been discoveries counteracting indefinite diversification which have been unifying and coordinating. Great generalizations such as Newton's laws of mechanics, Willard Gibbs' phase rule, and Einstein's relativity principle have made it possible to survey the ever widening fields of knowledge from central vantage points.

Since the last of the nineteenth century, physical chemistry has exercised a similar effect, though perhaps on a lower altitude of generality as a joint discipline. Developing rapidly side by side with the enormous proliferation of descriptive organic chemistry, physical chemistry first served to simplify and clarify inorganic and analytical chemistry. It is now, by peaceful penetration, performing a similar function in organic chemistry, including biological and physiological chemistry.

As a branch of physical chemistry, colloid science, developing still more recently, and concerned with the phenomena of structures having a characteristic range of dimensions and degree of organization, is also integrating in its overall effect. Colloid science clarifies and illuminates the physics and chemistry of a large fraction of the universe to which we are most intimately and continuously exposed. Most interesting

of all is the fact that the phenomena characteristic of living matter seem to operate to an important extent within the confines of this science.

Professor McBain has been a pioneer for many years in the exploration and exploitation of colloid science. It is understandable, therefore, that he should be able to present clearly, and with authority, the fundamentals of this science, and also to discuss with insight and discrimination its most recent advances and the unsolved problems that still confront it.

It is, therefore, with not unexpected profit and pleasure that I have read this present volume.

<div align="right">ARTHUR B. LAMB</div>

PREFACE

This book is designed to give a clear picture of Colloid Science, bringing out not only its immense scope but also the main regularities and generalizations that have been established within it. Colloid Science is a relatively modern subdivision of human knowledge, although its existence was clearly recognized in 1861 by Thomas Graham, who not only first appreciated its importance and generality but also coined the name "Colloid" (from the Greek word for glue) to represent this newly recognized variety of matter.

Today it is realized that most of the materials of our ordinary experience are colloids and that phenomena characteristic of colloids are all but universal in the physics and chemistry of both the organic and inorganic world, the living and the dead. Nevertheless, the term colloid is unfamiliar to the educated layman and indeed to many scientists, as contrasted with such recently publicized terms as vitamins and atomic energy. There is no mistaking the importance of colloid science in industry; and it should not be overlooked in academic and lay circles.

The approach in this book is that of a physical chemist convinced that the general laws of physics and chemistry prevail throughout this subject, that superimposed upon them are further laws of colloid science, and that all these are operative in living organisms, governing much of their complex and fascinating behavior.

The reader of this book should remember that each general statement has innumerable practical applications, which, if taken seriously enough, can be turned to great practical account. No branch of science offers more varied and striking behavior or more intriguing experimental demonstrations.

<div align="right">

J. W. McBain

</div>

CONTENTS

COLLOID SCIENCE

INTRODUCTION

The word "colloid," even in the relatively brief period since Thomas Graham proposed it, has acquired a meaning far wider and more significant than he had remotely imagined.* It now covers innumerable materials as different from glue, the equivalent of the Greek "κόλλα," as are Purple of Cassius and such everyday colloids as soap solutions, rayon, and bakelite. In fact, colloids constitute most of the materials of ordinary life.

The broadest definition of colloid science is the study of non-crystalline matter having an organization intermediate between that of independent molecules and of living matter, only limited by the proviso that this definition does not include more amorphous or gaseous matter.

Some have attempted to define colloids as a state of matter where the physics and chemistry of surfaces are dominant; that is, where the surface properties outweigh the bulk properties of the material. A little consideration, however, shows the inadequacy of such a definition. Not only does it ignore the marvelous organization which colloids such as soaps and detergents exhibit, as described in Chapter 17 and later chapters of this book, but it also does not describe such characteristic phenomena as the scattering of light, the mechanical and flow properties which interest the engineer and many an industrialist, or the properties of texture and touch which are of such importance in practical life and technology. It misses, too, such a remarkable phenomenon as the flakes of cetyl sulfonic acid flexing and swimming through a solution as described and shown in photographs by

* For general and text references throughout the book see the list at the end of each chapter.

Darke, McBain, and Salmon in 1921 in the *Proceedings of the Royal Society*.

Others have attempted to define colloids merely by the stipulation that they are composed of particles having at least one dimension lying within the range of ten Ångströms and one micron, i.e., 0.000,000,1 and 0.000,1 cm, corresponding respectively to about ten times the diameter of an oxygen atom and the diameter of a particle barely visible in a high-power microscope. This is, however, an arbitrary and artificial definition which also neglects entirely the important organization that we now know colloidal matter to possess.

The organization characteristic of colloids can be described by saying that colloidal matter consists of three coordinating factors: (1) particles, (2) a continuous medium, and (3) a stabilizing agent. These factors may be further briefly described individually as follows:

(1) The particles which are the distinctive units of colloids are not in general single molecules, but commonly each consists of a large number of molecules. The particles may be crystals or droplets or bubbles. They may be wholly separate from each other and, therefore, independent; or more usually they may be grouped to form still larger structures, some of which may have a high degree of organization, whereas others may be merely loose aggregates. In all cases the material acquires new and characteristic properties.

(2) The continuous medium may be anything -- a gas, a liquid, or a solid. Its continuity is broken only by the above-mentioned particles themselves or by their points or surfaces of attachment to each other. If the particles are sufficiently close, as in rubber or glass or paper, the continuous medium may be almost eliminated.

(3) The stabilizing agent must be of a dual nature, having affinity to both the particle and the medium which are united by it. In certain important cases, it may be supplied by polar chemical groups on the surface of the particle itself.

Sugar is soluble in water because of the numerous hydrogen atoms and, more especially, hydroxyl groups which it presents to a solvent consisting of those elements. However, these simple molecules maintain their independence even in the most concentrated aqueous solutions.[1] These solutions are termed crystalloidal solutions, or true solutions, in which the molecule or ion is the unit.

Some of the most important colloids are likewise spontaneously soluble. Their solutions are stable in the same strict thermodynamic sense in which solutions of sugar or salt are stable. These colloids are those whose particles are self-stabilizing, whose molecules organize

themselves into particles. At least part of each molecule must have an affinity for the solvent, although another part may be insoluble. In such truly stable colloidal particles, the molecules must so arrange themselves that some soluble part is exposed to the solvent, and thus they provide their own stabilizing agent. A cross section through such a stable colloid is given in the diagram in Figure 1·1, where the insoluble hydrocarbon chains are turned inward, in a neutral particle of soap, and the heavily hydrated carboxylate group is presented to the surrounding water.
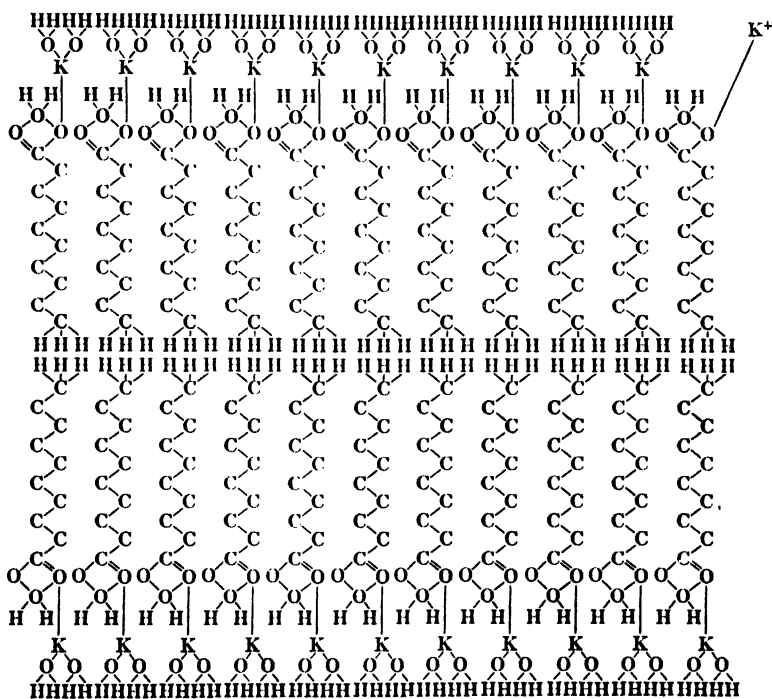
FIGURE 1·1 Diagrammatic cross section of a stable colloidal particle illustrating the principle of "like to like."

Solutions of such intrinsically stable colloids are readily prepared by mere contact with the appropriate solvent. Thus, proteins and gelatin or glue dissolve in water, nitrocellulose in nitroglycerin or acetone, rubber in benzene. On the other hand, to produce a "solution" of rubber in water, a stabilizing agent such as a water-soluble protein or a detergent must be present, as in latex or synthetic latex. (See Chap. 23.)

Very many colloidal particles are of the unstable type, depending for even temporary existence upon a chance stabilizing agent. Such

an example is pictured in Figure 1·2, which shows a cross section of a minute, true crystal of platinum metal whose surface is covered and stabilized by hexahydroxyplatinic acid.[2] Here again the outside of this particle is shown as presenting only hydrogen and hydroxyl to the surrounding water. The properties of such a particle depend very little upon the nature of the material in the interior as long as it suffices to anchor the stabilizing agent. Hence, we have adopted for the unstable colloids the slogan of the paint manufacturer, "Save the surface and you save all."

$$
\begin{array}{cc}
\mathrm{HO}\diagdown\ \underset{|}{\overset{\mathrm{H}\quad\mathrm{H}}{}}\ \diagup\mathrm{OH} & \mathrm{HO}\diagdown\ \underset{|}{\overset{\mathrm{H}\quad\mathrm{H}}{}}\ \diagup\mathrm{OH} \\
\mathrm{HO-Pt-OH} & \mathrm{HO-Pt-OH} \\
\mathrm{HO}\diagup\quad\diagdown\mathrm{OH} & \mathrm{HO}\diagup\quad\diagdown\mathrm{OH}
\end{array}
$$
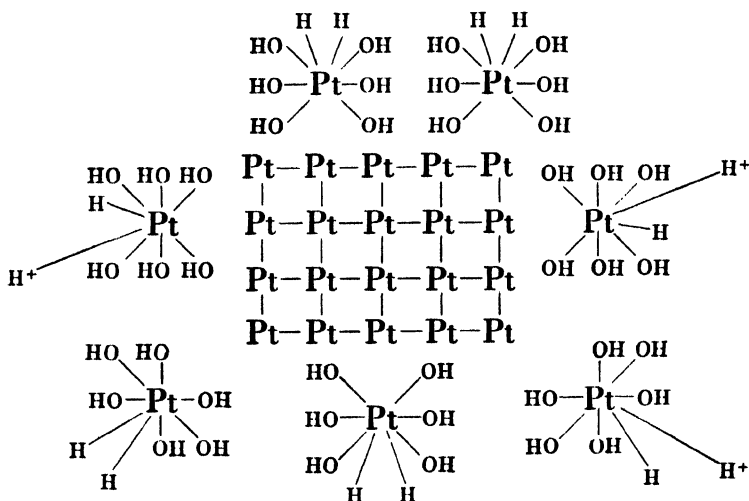
FIGURE 1·2 Diagrammatic cross section of an unstable colloidal particle consisting of a true crystal of platinum stabilized by a coating of soluble hexahydroxyplatinic acid.

In some cases, so large an amount of effective stabilizing agent is added that colloidal particles which otherwise would be unstable are given a high degree of stability for practical purposes. This is done consciously with some commercial preparations of metals for medicinal use, such as Paal's metal sols, that may be evaporated to dryness and dissolved again in water. It has been unconsciously done in Odén's sulfur. As a rule, when it is left to chance, the results are erratic and fleeting.

Recently it has been shown that a stable, protective colloid may impart genuine stability to otherwise insoluble material so that this even dissolves spontaneously and participates in true reversible equilibria.[3] This phenomenon is called solubilization.

A characteristic feature of colloids is the association that they exhibit. It acts on the principle of "like to like," and is, therefore, of

the cohesive, van der Waals' type of chemical attraction which, for example, is operative in forming double molecules in the vapor of acetic acid or in producing the ordered array of a molecular crystal. Association is manifested alike in the formation and in the properties of colloids. Without it, there would be no films, fibers, glasses, resins, nor plastics; no jellies, gels, liquid crystals, coagula, nor gelatinous precipitates; no high viscosity (Graham's colloidoscope); no opalescence; no electro-osmosis nor streaming potential; no membrane equilibrium nor any of the effects defined by Loeb as "colloidal behavior"; no stabilizing agents nor protective action or sensitization; no stability however transient of the great majority of colloidal particles which are dependent upon structural arrangement; not even the folding of the long chain molecules that gives the characteristic properties to such colloids as wool or rubber; none of the beautiful and varied architecture that utilizes as building stones the colloidal particles, which are themselves products of association.

It clearly emerges from the foregoing discussion that another salient feature of colloidal matter is the enormous development of surface relative to the amount of matter present. A familiar calculation shows that if 1 cc of material were subdivided into a thousand million million cubes each 10×10^{-8} cm or 10 Å or 1 mμ in diameter, the area of this 1 cc would be 6000 square meters, or nearly 1.5 acres. It is no wonder that the properties of colloidal particles are so largely those of their surface. It is not surprising that these colloidal properties so often outweigh the purely chemical ones of the interior material of the particles; particularly if, in referring to the colloid, we forget, as we often do, to mention the stabilizing agent altogether and think only of the material in Figure 1·2 as "platinum" and again forget that the particles are so often built up into larger formations that confer upon them additional properties.

In the diagrams of Figures 1·1 and 1·2, we have not yet called attention to the obvious fact that the colloidal particles are shown as carrying an electrical charge. The stability of the colloidal particles, that is, their failure to coalesce during the frequent collisions that are constantly occurring, is due to the physical interposition of the stabilizing agent which prevents actual contact between the uncoated interiors of the particles.

Many writers attribute a prominent role in stabilizing a particle to the electrical charge itself. It remains to be seen whether this is a primary or merely a concomitant factor. It is certain that diffuse ionic layers repel each other.[4]

It is a remarkable and a significant fact that almost all — possibly all — of the characteristic behaviors of colloids occur equally clearly in nonionizing solvents where no free ions are present. The subject of colloids is just as important in aqueous as in nonaqueous solutions. In nonionizing solvents one finds peptization, coagulation, mutual coagulation, jelly formation, thixotropy, and the other characteristic behaviors. Indeed some substances, such as quartz, feldspar, and fluorspar, which do not show thixotropy in water, do so in benzene. All these terms will be defined and discussed later on.

It is customary to refer to colloidal material as constituting a state of matter. That this is not a completely happy description is shown by the fact that the colloid condition has been always ignored by the numerous authors who variously describe a "fourth state of matter." It is more evidently a state of organization, comprising a number of organized forms of matter, which begins with the colloidal particle itself.

The word "colloid" was proposed by Thomas Graham (1861) in contradistinction to crystalline or crystalloidal matter, using the Greek "κόλλα," meaning glue. However, we have already noted that there are innumerable substances in the colloidal state which are not such complicated organic molecules as glue, starch, or hemoglobin. Thus, we might have mentioned with colloidal platinum other elements such as sulfur, gold, selenium, or sodium, or simple compounds such as silicic acid or silver chloride. All substances, like sodium chloride, may exist either in the crystalloidal form as simple ions or molecules or in the colloidal form. In the particular instance of sodium chloride the crystalloidal form is incomparably more stable. In soaps and other colloidal electrolytes the colloidal particles form spontaneously.

Soaps are crystalloidal in alcohol, forming only molecules and ions. In aqueous alcohol or water, however, colloidal particles are formed in equilibrium with ions. These are the colloidal electrolytes (Chap. 17). There are countless thousands of familiar examples of such colloidal electrolytes, which include most dyes and nearly all ionizable substances of high molecular weight. Colloidal electrolytes have sometimes been regarded as transitional materials between crystalloids and colloids, but that conception applies rather to the transition between the dextrins and the sugars where, as the size of the dextrin particle gets less and less, the distinction between particle and molecule, or double molecules, is lost.

Colloidal particles are often best distinguished empirically from ordinary molecules by the fact that they cannot enter the fine interstices of sufficiently compact colloidal jellies or membranes which

allow ordinary molecules and ions to pass freely — diffuse or filter through.

Colloidal particles are often revealed by the visible opalescence to which they give rise. Usually they are too small to be resolved in the microscope, and the colloidal solutions or sols will pass through ordinary filter paper, because they are less than 10,000 Å in at least one dimension.

Unstable colloidal particles primarily consist of materials which are insoluble in the surrounding medium. Such unstable particles are formed, for example, when mercuric sulfide results from the interaction of aqueous mercuric cyanide and hydrogen sulfide or when arsenic trisulfide is similarly produced from aqueous arsenious oxide.

Confusion between colloidal particles and true molecules or ions arises only if one attempts to rest the distinction upon an arbitrary basis of size alone. Some fibers and films have only two or even only one of their dimensions within the range of 10 to 10,000 Å, the others being far greater. Most ions and molecules of classical chemistry are from 3 to 10 Å in diameter. The conception of the simple chemical molecule or ion is well defined; the atoms are held together by primary valences. "A complex molecule is a group of covalently linked atoms, in which some at least of the shared pairs of electrons are not drawn one from each of the two atoms they link."[5] Two molecules associating together are conventionally and conveniently referred to as a double molecule or dimer, as in the case of acetic acid in vapor or in benzene. However, almost always when association is operative at all it proceeds far enough to build up a particle containing a large number of molecules.

Later we shall set up the generalization that with respect to a number of properties, single, independent, compact colloidal particles behave like single molecules or single ions, differing only in weight or size. Such properties include light scattering, osmotic effect, and characteristics which depend upon the kinetic energy of the particle, molecule, or ion.

When we turn to greatly elongated molecules, such as those of cellulose or proteins, where there is a very long chain of primary valences, we find that as long as these extensive molecules remain independent of each other and do not associate, their properties are in many respects those of ordinary single molecules. However, the reactive groups are often spaced so far apart, compared with the range of direct molecular attraction, that these reactive groups are largely independent. Then the reactions are like those of groups attached to

an ordinary surface or like those of a group of separate small molecules. Furthermore, the different segments of the long chain molecule may vibrate or move semi-independently, like the beads on a necklace; whereas the ordinary molecule is a kinetic unit.

We must emphasize that the colloidal condition may comprise any substance and any medium. Although great attention is given to colloidal solutions or sols, these are often not the most important forms, in comparison with jellies, gels, curds, precipitates, coagula, films, fibers, membranes, etc., as is evident in such industries as ceramics, plastics, textiles, paper, cement, leather, or rubber.

Wolfgang Ostwald in a famous passage [6] taken from a book written in the trenches, which also includes a tribute to the authorities on colloids belonging to other nations with whom his own nation was then at war, writes as follows of the universality of the colloidal condition:

> The clothes you wear, be they wool, cotton or silk, are animal or plant gels. They are dyed with colors which, in many instances, as the indigos and the blacks, are colloid in type. In the process of dyeing, adsorption and other colloid-chemical reactions take place between the colloid substrates of the fabric and the colloid dyes which color them. The leather of your shoes is an animal gel, closely related in its general properties to that prototype of the colloids, gelatin. Leather is tanned with substances of which the majority are colloids, and the whole process of tanning is punctuated with the colloid phenomena of hydration, dehydration and adsorption. The wood of the chairs in which you rest is made of cellulose, which in all its various forms is colloid in nature. The colloid swelling of wood, as I emphasized earlier, was used by the old Egyptians to aid their quarrying of stone. The woods of your chairs are held together by glue or with metals. You already know glue to be a colloid, but it may surprise you to learn that colloid chemistry has much to say in metallurgy and that steel, for instance, is a colloid solid solution. We shall return to this question. The paper upon which you write is essentially cellulose, in other words, again colloid. It has been given a body by being mixed with water-glass, with rosin or some similar material, in other words, with various colloids. The ink in your fountain pens is probably also colloid if it is the ordinary iron tannate, and colloid, too, is the hard rubber of your pen holders, prepared from that notoriously colloid mother substance, soft rubber.
>
> I could continue this list indefinitely, pointing in this manner to one colloid after another in your immediate surroundings and belonging to the things of your every-day life. Perhaps you think — perhaps since yesterday's lecture you think you know — that I am possessed of a colloid mania because I see colloids everywhere. Let me admit that I do see colloids everywhere, only I do not believe that because of this I must be adjudged insane. It is simply a fact that *colloids constitute the most universal and the commonest of all things we know*. We need only to look at the sky, at the earth, or at ourselves to discover colloids or substances closely allied to them. We begin the day with a colloid practice — that of washing — and we may end it with one in a bedtime drink of colloid tea or coffee. Even if you make it beer, you still consume colloid.

Sir William Bragg [7] in a presidential address to the Royal Society has pointed out the key position that colloidal magnitudes now present

in all lines of research — medical, industrial, metallurgical, and purely scientific. They are too small for the microscope and too large for usual x-ray methods. It is of ever increasing importance not only to develop new and definite techniques for making these colloid regions accessible to detailed examination but also to apply to the fullest extent and in quantitative manner the materials and methods that have already been introduced. A most promising approach lies in further development of the electron microscope.

The industries, through the ages and at the present time, have accumulated most of the actual information on various materials in the colloidal condition and the rate of this accumulation has of late been greatly accelerated, although publication has less often kept pace with it. Since it is only within the present generation that colloid science has become a recognized discipline, it is not surprising that the scientific worker has been hard pressed to interpret and correlate the multitude of specific observations. The enormous practical importance of such materials as the plastics renders valuable every lead that theory can give.

Historically, with respect to pure science, it is only necessary to mention the "potable" or colloidal gold of the alchemists and to refer to the efforts of Thomas Graham to distinguish and characterize the colloidal condition. To him we owe that apparently necessary means of scientific propaganda, a system of new names and terminology comprising such terms as sols for colloidal solutions, gels for jellies and various gelatinous coagula, peptization, pektization, colloids, crystalloids, and many others which long remained outside the use of other men of science.

No field of science has progressed more during the past generation. The psychological turning point was the demonstration by Siedentopf and Zsigmondy [8] in 1903 that particles of certain gold sols were definitely revealed in the ultramicroscope. This settled a question of long standing and opened the way to further definite experiment. Zsigmondy devoted the remainder of his life to the colloid field.

The quantitative study of colloids by unambiguous methods had proved singularly elusive, and it is interesting to note that four of those who provided a definite framework for investigation have been awarded the Nobel prize. The first was Zsigmondy.

Einstein resolved the apparent miracle of perpetual motion of the particles in the ultramicroscope by the simple postulate that all particles, including molecules, have the same mean kinetic energy and therefore the slow motion of the larger particles merely reveals the

thermal energy of the molecules, in verification of the kinetic theory of heat. Einstein[9] was also responsible for the quantitative formulation of the simplest case of diffusion, and also of viscosity, based upon the ordinary laws of hydrodynamics.

Svedberg by his remarkable invention of the ultracentrifuge devised a method of determining particle size and molecular weight even in the most difficult colloidal sols.

Langmuir provided clear-cut conceptions of adsorption and surface films and experimental methods for their investigation.

Apart from the sustained work of these and many other brilliant leaders, we need mention only four more names: Wolfgang Ostwald for his editorial organization of the whole subject, to which he has devoted the *Kolloid Zeitschrift* (1906) and the *Kolloid Beihefte* (1909-1943); von Weimarn for his enthusiastic experimental study of so many colloidal materials; Freundlich for his sound physical-chemical judgment and encyclopedic knowledge; and lastly Bechhold for his development of yet another quantitative tool, the ultrafilter.

A group of papers of exceptional historical interest published in the nineteenth century has been assembled by the Colloids Committee of the British Association and edited by Hatschek.[10]

Finally we may attempt to assess the value of the definite conceptions of colloidal particles recorded diagrammatically in Figures 1·1 and 1·2 by pointing out the large number of properties that may be deduced or surmised from a study of these diagrams. We learn one of the origins of the electrical charges on colloidal matter and readily predict the importance of adsorption of all kinds, including the reactions of colloids with each other. We therefore see how to augment or decrease or reverse the electrical charge, and we note the concurrent effects upon stability. We see how primary particles may adhere to each other but yet retain their individuality, forming larger structures or even elastic jellies. From the charges and from the fact that the particles can never wander far from compensating ions of opposite charge, we can deduce not only their migration in an electrical field but also the amount of conductivity which they and their accompanying ions will exhibit. Using Einstein's dictum, we estimate the osmotic effect or the diffusion of a given number of particles and can readily explain the influence of electrolytes and their distribution upon these properties. Lastly, from the diameter and nature of the particles, we see whether or not they will pass through colloidal sieves or ultrafilters of given size, and in the simplest cases we can predict the amount and color of scattered light.

We shall find that the electrical charges produce the innumerable colloidal electrolytes, many of whose properties are shared in lesser degree by all charged colloids. Another important finding is that charges of particles are surrounded by a diffuse ionic cloud, such ionic atmospheres resulting in a repulsion between charged surfaces and particles extending over a range of hundreds or thousands of Ångström units.

It is the purpose of the following chapters to deal with these and many other aspects of the colloidal condition. A number of colloidal materials are of such importance and interest, and some exhibit such characteristic features that separate chapters are devoted to them.

REFERENCES

Text References

1. MCBAIN, J. W. and STUEWER, R. F., *J. Phys. Chem.*, **40**, 1157 (1936)
2. PENNYCUICK, S. W., *J. Chem. Soc.*, 1928, 2108; *Kolloid Z.*, **49**, 407 (1929)
3. MCBAIN, J. W. and MCBAIN, M. E. L., *J. Am. Chem. Soc.*, **58**, 2610 (1936)
4. VERWEY, E. J. W. and OVERBEEK, J. T. G., *Theory of the Stability of Lyophilic Colloids*, Elsevier Publishing Co., Inc., New York (1948); BRADY, A. P., *J. Phys. Coll. Chem.* **53**, 56 (1949)
5. SIDGWICK, N. V., *J. Chem. Soc.*, 1941, 433
6. Quotation reprinted by permission from *Introduction to Theoretical and Applied Colloid Chemistry*, by W. OSTWALD and M. H. FISCHER, published by John Wiley & Sons, Inc., New York (1917) p. 179
7. BRAGG, SIR WILLIAM, *Proc. Roy. Soc.*, **163A**, 472 (1937); see also *Nature*, **140**, 954 (1937)
8. SIEDENTOPF, H. and ZSIGMONDY, R., *Ann. Physik*, **10** (4), 1 (1903). According to C. V. Raman (*Phil. Mag.*, **17** (6), 495 (1909)), G. Dubern (*Indian Eng.*, April 7, 14, and 21, 1888) anticipated the ultramicroscope in the form of the simple rhomb for sideways illumination subsequently employed in 1906 by Cotton.
9. EINSTEIN, A., *Ann. Physik*, **19** (4), 289 (1906); **34**, 591 (1911); *Z. Elektrochem.*, **14**, 235 (1908)
10. *The Foundations of Colloid Chemistry, A Selection of Early Papers Bearing on the Subject*, Ernest Benn Limited, London (1925)

EMULSIONS AND FOAM

EMULSIONS

A very large literature has grown up around the subject of emulsions on account of their vital importance in a surprisingly large number of industries and in medicine. The best monograph on the subject is that of Clayton [1] and the most readable introduction, that of Sutheim.[2] An illuminating symposium was held in London in 1934, comprising papers whose subjects ranged from mechanism and patents to toxemia, food, insecticides, leather, rubber, and bituminous asphalt; this symposium has been reprinted, supplemented by a few new contributions.[3] Very readable, short, general papers are those of Holmes,[4] Thomas,[5] and Neville.[6] Further discussion of this subject is to be found in Chapter 17.

Emulsions are discussed at this stage for the purpose of focusing attention upon the importance of the specific properties of the stabilizing or emulsifying agent that is invariably present. It presents a close analogy to the stabilizing agent already emphasized as a feature of colloids.[6]

An emulsion consists of droplets of one liquid suspended in another continuous liquid. Such an arrangement would hardly be more stable than an effervescing liquid were it not for the indispensable presence of a stabilizing agent in the boundary or interface between the two liquids. The essential properties of an emulsion are largely determined by this agent. It may even be possible for an emulsion to be stable indefinitely if the droplets are of just such a size that the total interface exactly accommodates a rather insoluble stabilizing agent.[7] Indeed, the work of Schulman and his collaborators [8] and of Winsor [9] indicates that an insoluble amphiphilic substance like dodecylalcohol in the presence of a detergent insists upon a definite minimum area of inter-

face for each of its molecules, a definite "*Lebensraum.*" Such systems form a continuous bridge between solubilization, discussed in Chapter 17, and emulsions, as described in the present chapter.

Many of the properties of emulsions are common to colloids because in both cases they are so largely those of the stabilizing agent. However, it must be pointed out that very few colloidal particles are liquid. It is a relic of older thought that certain colloids are sometimes termed "emulsoids." As a rule, emulsions are very much coarser than colloids and may be examined with a microscope or sometimes even with a lens. They serve as models of colloidal suspensions or unstable colloids, such as "gold," rather than of the so-called "emulsoids" or self-stabilizing, truly stable colloids, such as soaps or proteins.

The Two Types of Emulsions

Emulsions are commonly classified as water-in-oil and oil-in-water. These are given the technical designations *w*-in-*o* and *o*-in-*w*, respectively, or *w*/*o* and *o*/*w*, the first mentioned constituent referring to the globules. A perhaps clearer terminology is oily emulsion of water and aqueous emulsion of oil. An example of the former is butter, and of the latter some margarines. Cold creams and "skin foods" are usually water-in-oil, whereas day creams, vanishing creams, and complexion milks are usually oil-in-water. Figure 2·1 a and b illustrates the natural occurrence of water-in-oil emulsions in crude petroleum. In emulsions the droplets are referred to as the disperse or dispersed phase and the second liquid as the dispersion medium. Other terms are internal and external phase, discontinuous and continuous liquid, and closed phase.
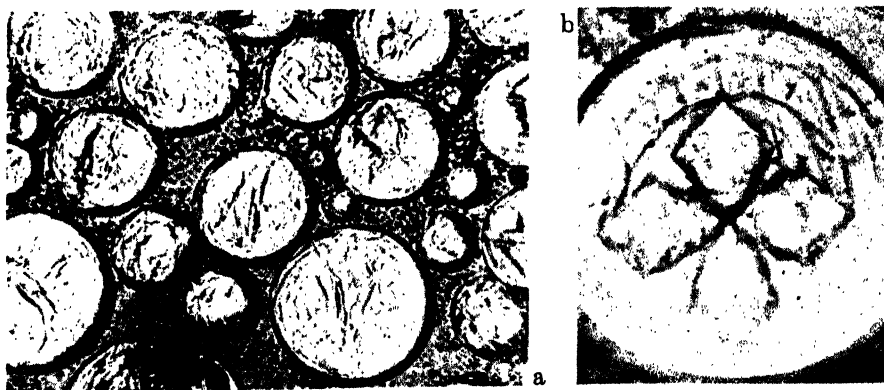
FIGURE 2·1 Natural petroleum from Fitts Pool, Oklahoma. Brine emulsified in oil, showing that some emulsions can last indefinitely. Note the wrinkled skin around the droplets and the crystals of salt in the saturated brine. Magnification of original photograph: a, 200 ×; b, 1200 ×.

A stabilizing agent is a "solution-link" (Thomas), and a link must be connected at both ends. Thus, a stabilizing agent must serve two masters. It must have an affinity for each of the liquids in the emulsion, and it must exhibit a distinct preference for one, but not such an overwhelming preference as to leave it indifferent to the other. Even if the stabilizing agent is a solid, these statements hold true.

Test for Type

The first question one would ask in examining an emulsion is whether it is an aqueous emulsion of oil or an oily emulsion of water. The first method of testing is that of Brailsford-Robertson and consists of examining a spread-out film with a microscope, the specimen having been sprinkled with a powdered, oil-soluble dye such as red Sudan III. The oil will be present either as bright red circles in an undyed background or as a continuous red film interspersed with undyed circles. In the first case this is obviously an oil-in-water and in the second a water-in-oil. The second method is that of Briggs, who tests the emulsion for spreading upon each of two liquids separately. Upon one it spreads by mere dilution or extension of the continuous phase; it does not spread upon the other. A variant is to place it upon filter paper to see whether the latter receives a translucent oil stain from the continu-
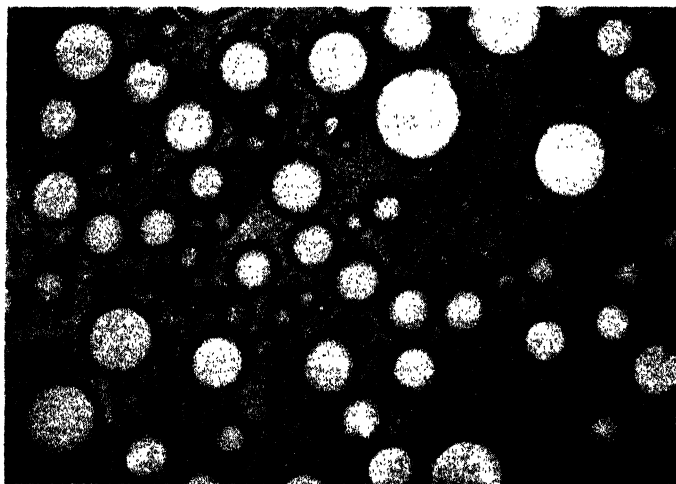
FIGURE 2·2 An emulsion of bituminous asphalt in water. Modern highway construction is largely based on the use of these aqueous emulsions; 4,000,000 tons of bituminous asphalt or petroleum pitch are used annually in this form in the United States. From Garner, F. H., Gabriel, L. G., and Prentice, H. J., *Modern Road Emulsions*, The Road Emulsion and Cold Bituminous Roads Association, Ltd., London (1934).

ous phase or is untouched by oil globules. Another variant is that of Pickering, who pointed out that the emulsion will mix perfectly with more of the continuous liquid. Again, Newman [10] placed a drop of pure liquid in an emulsion near the wall of the vessel where it could be seen; if it were the same as the continuous medium, it merely dissipated and diluted the emulsion; if not, it remained clearly delineated and either rose or sank depending on its density.

A third method suggested by Clayton is to test the conductivity either with direct current or with alternating current and a loud speaker or a lamp, because the conductivity of a continuous aqueous liquid is of a higher order of magnitude than that of oil. A fourth method was suggested by Joshi, who observed a striking change in viscosity (or surface tension) when o/w changed to w/o, or vice versa. Recent workers have noted a sudden, high maximum viscosity during this reversal of phases, owing to the complex or multiple nature of the emulsion, where large drops of one liquid may be filled with smaller drops of the other, and these may also enclose still smaller drops of the first, leaving a much smaller portion of either liquid as a continuous medium. Such
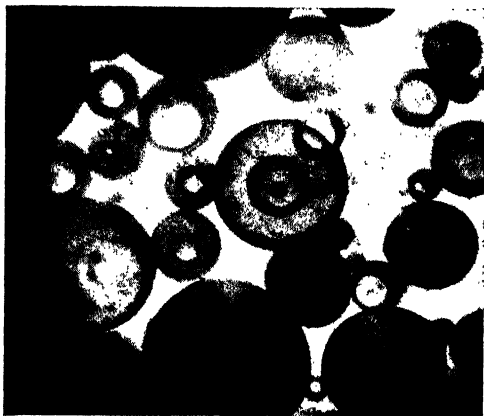
FIGURE 2·3 Photomicrograph of multiple emulsion. In center is a drop of a quinqui-multiple emulsion. Taken from Seifriz, W., *J. Phys. Chem.*, 29, 745 (1925).

a multiple emulsion is illustrated in Figure 2·3. In the center of Figure 2·3 is shown a photomicrograph of a drop of quinqui-multiple emulsion, in which three oil and two water globules alternate one within the other.

The fifth method [11] rests upon the fact that droplets of refractive index differing from that of the continuous liquid act as lenses, and only one side appears illuminated. A sixth method is that of Hauser which uses a fluorescent microscope illuminated only with ultraviolet light; if the droplets of the medium are not naturally fluorescent, a fluorescent material, such as anthracene, is added. A last miscellaneous group comprises methods of the expert familiar with his product who recognizes its typical properties from some casual indication such as its appearance, the sound of the splash when it falls into a vessel, or, if it is a stiff emulsion, its behavior on being cut.

Preparation of Emulsions

The task is to subdivide one liquid into the other and to stabilize the globules formed so that they do not coalesce. Much specific experience is often necessary to obtain optimum results because it is easy for unsuitable mechanical treatment to subdivide the wrong liquid or to induce coalescence by damaging the stabilizing film. A striking example of the latter is the lecture experiment of Briggs [12] in which 80 cc of benzene is shaken with 20 cc of water containing 1% of sodium oleate. One vessel is shaken violently for 3 minutes while the other is given only seven single shakes, each of 30 seconds duration, and allowed to rest undisturbed in between. The latter is wholly emulsified at the end of this time, while the former largely separates into its original constituents.

As a rule, the liquid to be dispersed is added in small quantities to the continuous phase whereupon it is spread into a thin, unstable film which spontaneously breaks up into droplets under the influence of surface tension. (See Fig. 2·4.) The formation of such a film may be effected by bubbling, boiling, spraying, or, more frequently, by grinding or by forcing the mixture through a narrow interspace between a rapidly revolving rotor and a stationary wall, or merely by forcing it through a poppet valve against the pressure of the spring.

FIGURE 2·4 Bursting of a mercury coated bubble to yield a mercury emulsion with varying size of droplet. Taken from Kremnev, L., *Kolloid Z.*, 68, 20 (1934).

Many varieties of emulsifying machines are used, ranging from the homogenizers used in reducing the sizes of the globules in milk or cream to the colloid mills of the Plausen, Premier, Charlotte, Day, Hurrell, Eppenbach, and Travis [13] types, some of which combine cutting or beating with shearing action. Figure 2·5 illustrates the homogeneity produced by one such mill.

Ultrasonic waves [14] may be used to produce emulsions. Their action is due to cavitation in the presence of dissolved gas; weak waves have the opposite effect by collecting liquid in the nodes. It is frequently of great advantage to have the emulsifying agent formed in the interface by chemical reaction, as in the formation of soap from aqueous alkali and an oil containing fatty acid, or by the hydrolysis of various derivatives of the higher fatty acids. Emulsions are readily formed where
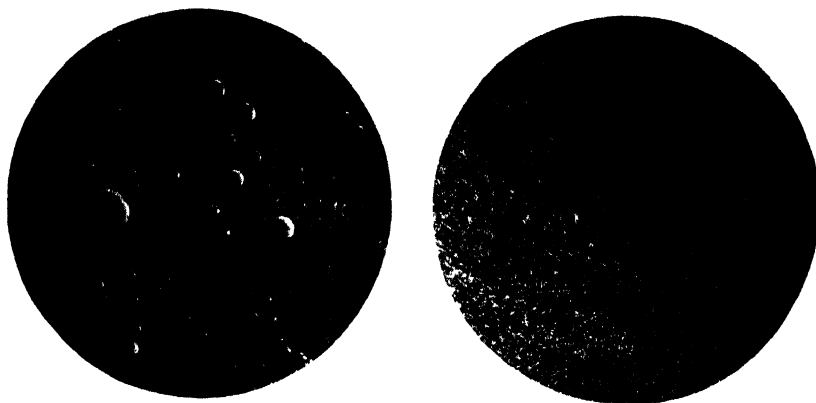
FIGURE 2-5 An emulsion before and after grinding in a colloid mill. Courtesy of Premier Mill Corporation, Geneva, New York.

agents are present which sufficiently reduce the interfacial tension or where the oil phase, previously dissolved in alcohol or other solvent miscible with water, is suddenly thrown into a large volume of the latter and thereby precipitated as insoluble fine globules that have no opportunity to coalesce before being stabilized [15] by the agents present or by the alcohol.

Spontaneous Emulsification

Emulsification can be completely spontaneous in strict absence of mechanical stirring, and even downwards against the direction of gravity for a light oil emulsifying in heavier water. This phenomenon was discovered in 1878 by Johannes Gad.[16] There are three distinct mechanisms and explanations.

The first mechanism, pointed out by Quincke [17] in 1879, is that lowering of interfacial tension by a solute is localized and causes such violent spreading that the resultant turbulence entraps globules of the upper liquid. This result accounts for part of the success of the experiments of Caldwell just referred to; however, if in his experiments a stabilizing agent is omitted, the emulsification is quite transient and is shortly followed by coalescence of the droplets. In such experiments as those of Gad, oleic acid is first dissolved in the oil phase, and sodium hydroxide in the water, and then when they are brought into contact, soap is formed in the interface, lowering the surface tension violently and remaining in the interface to protect and stabilize the droplets. This procedure is often adopted, and the emulsion is as much due to interaction as it is to the mechanical treatment given it.

The second mechanism is that suggested by Gurwitsch [18] in 1913 to explain spontaneous emulsification of petroleum oils containing naphthenic acids in aqueous alkali, namely, diffusion across an inter-face carrying solvent molecules along. Likewise, Raschevsky [19] in 1928 from purely theoretical considerations agreed that spontaneous emul-sification can be the result of diffusion processes alone. McBain and Woo,[20] who give references to the literature, observed that spontaneous emulsification may occur into either liquid or into both, depending upon the concentration of the diffusing substances. The presence or formation of colloids is quite unnecessary, for excellent examples of either behavior can be obtained with mixtures of alcohol, toluene, and water. It is evident that, as has been found by Hartley and by McBain, diffusing columns have a "push" effect driving along with them other substances by molecular collisions. In this way it is easily possible to carry an excess of any constituent into another region or liquid, con-trary to the natural direction of its movement, or overshooting the equilibrium, which is subsequently slowly restored.

FIGURE 2·6 Spontaneous slow streaming of diglycol laurate downwards into water, photo-graphed after one hour.

In spontaneous emulsification concentra-tions are important. There is often an upper and a lower limit for a particular effect or for the formation of either type of emulsion. If soap is being formed at the interface or is diffusing from one liquid to the other, emul-sification may be spontaneous even, as in some cases discussed in the next paragraph, if the soap is already formed and in its final solvent. Spontaneous emulsification is of great impor-tance in the use of "soluble" oils [21] and in the formation of bituminous asphalt emulsions for road making.

A third mechanism is required to account for the many cases in which a *pure* liquid such as xylene or styrene spontaneously sub-divides itself into innumerable small globules when placed upon an aqueous solution of a soap or detergent. Many such examples have been observed in the study of emulsion poly-merization, as for example, of synthetic rub-ber. The energy for such spontaneous emul-sification has been shown to come solely from the energy of sorption of the detergent within

FIGURE 2·7 A drop of diglycol laurate placed on a lens of water and spreading horizontally to give myelin forms.

the colloidal particles of the detergent and in the interface [22] between the two liquids.[23]

With somewhat higher concentrations the diffusion may produce curiously involuted and convoluted excrescences known as "myelin forms." [24] Figures 2·6 and 2·7 illustrate how diglycol laurate placed quietly on water streams downward and leads to emulsification and under different circumstances produces myelin forms.

Creaming and Proportions by Volume

If oil is suspended as independent drops in a heavier liquid such as water, the drops will tend to rise under the influence of gravity. This is known as creaming. A heavier oil will tend to settle. As the droplets move through the liquid they frequently sweep others along with them so that creaming is sometimes more effective in deep vessels than in shallow pans. In the cream the droplets are close together, and they may even adhere or clump together, but it is important to note that they have not coalesced. The latter constitutes the "breaking" of the emulsion.

Wolfgang Ostwald suggested that since 74% by volume represents the closest packing of spheres of uniform size,[25] this percentage might form an upper limit for the droplets of an emulsion, and an increase in percentage would be followed by a reversal of phases. This theory

was disproved by the famous emulsions of Pickering [26] in which 99 cc of paraffin oil were emulsified in 1 cc of aqueous potassium oleate, and Newman's similar emulsion of 99 cc of benzene in 1 cc of aqueous sodium oleate, and his emulsions of 96 cc and 94 cc of water in 4 cc of benzene or in 6 cc of carbon disulfide using magnesium or sodium oleates as agents. Such emulsions are so stiff and elastic that they may be heaped up to form a sharp cone. However, this proportion of 74% serves as a guide to the concentration of "cream" upon a dilute emulsion which has sedimented on standing without breaking. Likewise, it is sometimes related to the maximum concentration of a homogenized emulsion or to that where reversal of phases takes place if a suitable stabilizing agent for the opposite type is present.[27] It is, of course, the upper limit for uniform spherical solid particles in paint,[28] since distortion or reversal of solid phases would be impossible, but 52% is "critical" because of cubic close packing. Closer packing is possible if the spheres are of various sizes in the same emulsion.

Appearance

An emulsion is usually turbid, opaque, translucent, or opalescent, but if the refractive index of both liquids is made the same, it becomes transparent. Chromatic emulsions are produced when the refractive index of the two liquids is made identical for only one wave length of light.

Stabilizing Agents

(1) *Solids.* Pickering [29] and others found that numerous powders may stabilize emulsions, as in the case of insecticides. Basic salts, clay, lime, calcium carbonate, pyrites, and powdered glass are examples of oil-in-water stabilizers and carbon black of a water-in-oil stabilizer. Gelatinous aluminum hydroxide and the very effective polyvinyl alcohol have been recently introduced. The interface has a minimum area when the particles coat the outside of the droplet, as in Figures 2·8 and 2·9. The rule is that the liquid which wets the solid better, so that the interface meets the solid at an acute angle in that liquid, is the outer phase. If the contact angle is 0°, 90°, or 180°, no stabilization can occur because the solid will go wholly into one or the other liquid or exhibit no preference. In the intermediate angles lie the maxima for formation of emulsions.

A commercially important solid emulsifying agent may be taken for discussion. Gelatinous alumina [30] aged to become insoluble and
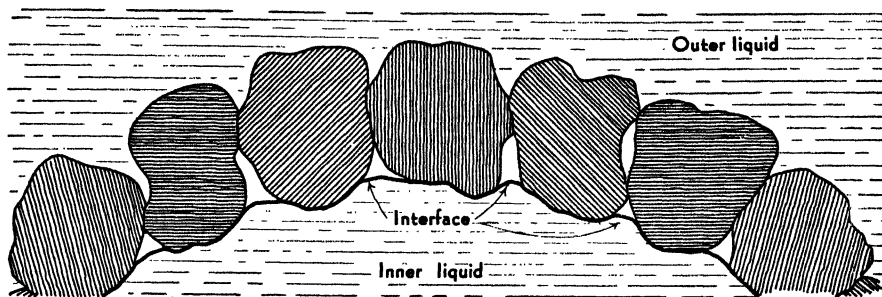
FIGURE 2·8 Particles of a finely divided solid acting as emulsifying agent. Taken from Thomas, Text Ref. No. 5.

FIGURE 2·9 Toluene emulsified in water by powdered pyrite. Through the courtesy of J. H. Hildebrand.

inert is a nonspecific emulsifying agent acting as a solid and not through reducing surface tension. While inert, it is edible, unaffected by molds or fungi, indifferent to moderate alterations of pH, and, if used in higher concentrations, acts as a suspending agent for solids as well as an anticreaming agent for the emulsion. It may even be used to treat other powders so as to make them readily wetted and suspended in water. The emulsion droplets may be made much smaller if, in addition, either sodium hydroxide or oleic acid is used.

(2) *Colloids and films of colloids.* Innumerable colloids have been used or patented for producing emulsions. Prominent are the ordinary and the heavy metal soaps, glue, casein, starch, agar, proteins, gums, egg yolk, dextrin, saponin, flour, milk, nitrocotton, resins, and the modern synthetic wetting agents. Some, like lecithin, are only partly hydrophilic, and like cholesterin, lanolin, the sterols, and the fatty acids are taken up by proteinaceous matter. In connection with the many cosmetic emulsions, both of w/o and o/w types, it may be recalled that the skin has been described [31] as a protein gel containing

about 27% protein, 2% partially hydrophilic lipoids or fatty bodies, and about 67% water. Some colloids may form tenacious, plastic, even visible films at the interface.[32] The formation of a tough film is the most effective method of stabilizing an emulsion or froth.

(3) *Polar substances.*[33] Most of the above are polar substances in which part of the molecule has an affinity for water and another part for oil. Holmes and Williams [34] have emphasized that the agent may not be colloidal, as with polar substances such as the alcohols, fatty acids, or benzonitrile. Other investigators have even used iodine to emulsify ether in water.[35] The alkali halides such as 0.5 N potassium iodide produce relatively stable emulsions of water in amyl alcohol although they increase the interfacial tension.

It is well to be reminded that the actual stabilizing agent in the interface is not necessarily identical either chemically or physically with the emulsifying agent or agents that were added to the bulk liquids. Such differences are rendered visible when a noticeable skin appears (see Figure 2·1) and are likewise made strikingly evident in experiments such as those of Deutsch [36] where various indicators and dyes are added to water covered by a layer of an indifferent liquid such as benzene, hexane, or carbon tetrachloride. When the system is shaken there is a vivid change in color due to the new species formed in the greatly enlarged interface. On standing, as soon as the layers have separated again, the color reverts to the original. Similar color changes are observed when such materials are sorbed on solid surfaces such as paper or white sand. Another instance where the agent as sorbed in the interface differs from that in the bulk liquid is soap, which often is sorbed in the form of acid soap.

(4) *Mixed stabilizing agents.* Again, a mixture of stabilizing agents may reveal their incompatibility, or, conversely, it may result in a strong reinforcement of the stabilizing action of either material alone.[37] Such favorable mixtures are ethyl alcohol and lycopodium powder, acetic acid and lampblack, dilute sodium oleate and clay, or a sulfated alcohol with sulfur.

Stability under a range of varying conditions may be insured by a suitable mixture of stabilizing agents, as, for example, one that is cation active with one that is anion active, such as a pyridinium derivative with a sulfated soap. A combination of an oil-soluble substance with a water-soluble substance can produce a very stable emulsion, as, for example, in the use of glyceryl monostearate with soap or the detergent known as sapamine. The investigations of Schulman, Rideal, and their collaborators [38] on the interaction between surface-

active agents are of general importance, not only for the study of emulsions but, more especially, for examining and explaining the behavior of membranes and their permeability and susceptibility to attack by various agents. As Rideal states, this is a new approach to biological problems such as lysis and hemolysis, sensitizing, agglutinating and immunizing reactions, lethal action, carcinogenic activity, penetration of drugs, bactericidal effects of detergents, and action of oestrogenic compounds. Some of their examples of mixtures tested as emulsifying agents for 10 cc of Nujol in 40 cc of water are given in Table 1, and the type of film obtained with cation-active emulsifying agents is shown in Table 2. Neither the oil-soluble cholesterol nor the

TABLE 1

Two-Component Films as Emulsifying Agents

10 cc Nujol containing 140 mg of Component A, with 40 cc water
containing 75 or 250 mg of Component B

Component A	Component B	Remarks
cholesterol	————	No emulsion
————	75 mg Na-cetyl-SO₄	Very poor emulsion
cholesterol	" "	Excellent emulsion (liquid)
oleyl alcohol (cis)	" "	Very poor emulsion
elaidyl alcohol (trans)	" "	Very good emulsion (viscous)
cetyl alcohol	" "	Very good emulsion (grease)
cetyl alcohol	250 mg Na stearate	Stability good
cetyl alcohol	" Na elaidate (trans)	Stability good
cetyl alcohol	" Na oleate (cis)	Stability poor
cholesterol	" Na oleate (cis)	Stability very poor
cholesterol	" Na elaidate	Stability good

TABLE 2

Positively Charged Emulsifying Agents and
Inversion of Emulsions

Emulsifying Agent	Interfacial Film	Emulsion Type
1. Lissolamine (pH 7)	None	None
2. (1) + cholesterol (pH 7)	Liquid + charge	Oil-in-water
3. (2) + Na₂SO₄	Liquid + charge	Oil-in-water
4. (2) + Na₃PO₄	Solid, neutral	Water-in-oil
5. (2) + tannic acid (pH 3)	Solid, neutral	Water-in-oil
6. Heptadecylamine (pH 7)	None	No emulsion
7. Heptadecylamine HCl(pH 2)	Liquid + charge	Oil-in-water
8. (7) + Na₃PO₄	Solid, neutral	Water-in-oil
9. (7) + Na₂SO₄	Solid, neutral	Water-in-oil
10. (7) + tannic acid (pH 2)	Liquid + charge	Oil-in-water
11. (7) + Na tannate (pH 7)	Solid, neutral	Water-in-oil
12. Heptadecylamine + 3N HCl	Solid, neutral (Multilayer)	Water-in-oil

water-soluble sodium cetyl sulfate is an effective emulsifying agent
alone, whereas the two together form stable films of emulsifying agent.
Cholesterol may be replaced by elaidyl alcohol, a trans-isomer, but
not by its cis-isomer. Similarly, the trans-isomer of sodium elaidate is
effective where the cis-isomer of sodium elaidate is not. The lissolamine
mentioned in Table 2 is cetyl trimethyl ammonium bromide, too
soluble in water, while heptyldecylamine is too soluble in oil, but a
mixture of them makes an excellent emulsifying agent.

Theory

A low interfacial tension or a high viscosity or the inevitable elec-
trical charges are each insufficient factors to stabilize an emulsion. A
solid coating, a polar compound, or a film is best. Sometimes it is evi-
dent that the stabilizing agent is present in the form of a tough, coher-
ent, or plastic film; this is noticeable in Figure 2·1 a.

Again it may be emphasized that the stabilizing agent in the inter-
face may not be the same as in the bulk solution. This is made visible
by color changes in the experiments of Deutsch [39] where water and
benzene are shaken together after the addition of colorless rhodamine.
When the extent of the interface is greatly magnified by the shaking,
the temporary emulsion appears bright pink. The color disappears
again as soon as the liquids separate. Again, soap is often sorbed in an
interface in the form of acid soap.

Bancroft made the empirical generalization that the continuous
medium is that in which the stabilizing agent is more soluble. Cassel [40]
has shown that Bancroft's rule unexpectedly rests upon a strict ther-
modynamic basis.

$$\ln p_r = \ln p_\infty \frac{2\sigma}{r(D_{internal} - D_{external})}$$

where the term on the left is the natural logarithm of the partial vapor
pressure of the stabilizing agent in the droplet of radius r, p_∞ is its
partial vapor pressure over the plane surface, σ is the interfacial ten-
sion, and D is the concentration in gm per cc in the droplet and the
continuous phase, respectively. McBain points out that it is obviously
much more difficult for two droplets to coalesce when immersed in a
solution of stabilizing agent than in the reverse case where they meet
in an indifferent medium and a slight rearrangement or resolution of
the molecules on their surfaces can lead to coalescence.

Many authors formerly held what is known as the oriented molecu-
lar wedge hypothesis which was based upon the hypotheses about the

relative dimensions of the head and tail ends of the soap molecule, the heavy metal soaps supposedly having several tails. However, the droplets are usually of a higher order of magnitude than that which would be so produced, and aluminum palmitate or stearate [41] in contact with water has been shown [42] to possess only one hydrocarbon chain or tail like a sodium soap, being a monobasic, not a tribasic, soap. Talmud and his collaborators discuss the effective hydration of the water-soluble group in its dependence upon the degree of close packing and orientation of the molecules of the agent and the effect upon the tensile strength of the film.[43]

Lewis [44] has discussed from the standpoint of interfacial film the distinction pointed out by Powis [45] between the adherence of droplets to form groups or clumps and the much slower further process of the coalescence of the individual adhering droplets to form one large spherical drop. Adherence or clustering gives enhanced and easily altered viscosity and a hindrance of settling that will be discussed under the term "thixotropy" in Chapter 10. The nature of the emulsifying agent may greatly modify the viscosity of emulsions containing the same proportion of the same oil in water.[46]

Reversal of Phases

Neutral soaps are insoluble in benzene but acid soaps are freely soluble therein. The higher sodium soaps are likewise nearly insoluble in cold water, as are all the acid soaps. Dilute soap solutions hydrolyze to form acid soaps, and less soluble soaps are readily salted out from aqueous solution by electrolytes. Presence of any heavy metal salt immediately produces a corresponding soap which is insoluble in water but soluble in benzene. It is evident that this complex of circumstances can yield complicated relations with emulsions that have often been misunderstood. Especially is this the case when soluble soaps like potassium laurate and sodium oleate are compared with insoluble soaps such as sodium stearate and when hydrolysis is not taken into account.

Bancroft's rule should be recalled: the phase in which the stabilizing agent is more soluble will be the continuous phase. Hence, if the solubility relations change with change of temperature, concentration, addition of salts, acid, or alkali, we may expect a reversal of phases to occur. This has frequently been observed. Even the volume relations may decide how much acid soap is formed in the benzene extract from the aqueous phase.

Bancroft was much influenced by the fact that ordinary soaps stabilize oil-in-water, whereas calcium and other heavy metal soaps stabilize water-in-oil. In Clowes' experiments, the emulsions were reversed by adding excess of one or other of these stabilizing agents. When a sodium soap stabilizes water-in-oil in the presence of much oil, it is due to hydrolysis of the soap forming acid soap and free fatty acid, which are both soluble in the oil but not in the water.

An analogous case is afforded by the Montgomerie type of bituminous emulsion. Here an asphalt containing a small amount of saponifiable material (naphthenic or asphaltogenic acid) is melted and stirred with an approximately N/20 solution of sodium hydroxide. The result is an asphalt-in-water emulsion which may be stable for years. However, since the sodium salt formed is sparingly soluble, an excess of alkali produces instead a thick emulsion of water-in-asphalt. Since the acid is weak, too dilute an alkali permits hydrolysis and again a water-in-asphalt emulsion results. The thin asphalt-in-water emulsion is instantaneously broken or reversed on contact with road building material or cement that supplies calcium or aluminum salts.

The mechanism of inversion is not entirely clear, but Figure 2·10, taken from G. M. Sutheim's *Introduction to Emulsions*,[47] suggests how the continuous phase becomes entrapped during the coalescence and inversion. A typical study is that of Rehbinder, Goldenberg, and Ab.[48]

FIGURE 2·10 Diagram showing how an emulsion reverses. Taken from Sutheim, "Introduction to Emulsions," through the courtesy of The Chemical Publishing Co., Inc., Brooklyn, New York.

Wellman and Tartar [49] observed phase inversion of benzene-water emulsions when stabilized with sodium stearate or palmitate. These soaps have very limited solubility at room temperature but within that range they form normal oil-in-water emulsions. When higher concentrations of soap are attempted, the rather oleophilic soap curd fibers remain undissolved and form the water-in-oil emulsion. On being heated above approximately 40° for palmitate and 55° for stearate, the soaps become freely soluble in water and the oil-in-water emulsion is re-formed. This reversal can be repeated in either direction by alternately bringing the soap into and out of solution. This illustrates the importance of knowing the properties of the emulsifying agent itself.

Stability of Emulsions

Mere permanence of an emulsion left completely undisturbed must be distinguished from stability under conditions of shaking, vibration, temperature change, centrifuging, or other stress. One emulsion that could be kept for months always separated into its two liquids when placed carefully on a glass slide; [50] others, particularly those with tough films, survive rough treatment without coalescence of the droplets. Stability is measured by the reciprocal of the initial rate of separation of the liquid from the internal phase in a centrifuge, [51] or change of area of emulsion interface, [52] or of size frequency. [53]

Breaking of Emulsions

Emulsions are sometimes extremely unwelcome in industrial processes, not only for their occasionally high viscosity but also where recovery of one of the pure liquids is desired. They may sometimes be broken by suitable mechanical treatment, by heating under pressure, by distillation, or by freezing. Chemical additions may destroy the stabilizing agent, as when strong acid is added to soap, or electrolytes may coagulate it. Excess of the internal phase may be added, or a mutual solvent for both phases, or a stabilizing agent of the opposite type. Frequently, a small addition of another competing agent, in itself an excellent emulsifier, such as soap, suffices to break the emulsion. Filtration through special filters such as Hatschek's magnesium carbonate powder has also been used. The charges on the droplet have likewise been turned to account in separating emulsions by electrolysis or treating with high tension electricity as in the Cottrell process. This method has been utilized industrially to some extent in breaking of emulsions of water in petroleum in California, the Gulf Coast, and in parts of South America, Russia, and Egypt.

A well-known example of the varied reagents that are used with troublesome water-in-oil emulsions of crude petroleum [54] is Barnickel's Tretolite consisting of Turkey red oils obtained by the reaction of sulfuric acid on fatty material such as castor oil. One drum completely separates ten thousand barrels of crude petroleum, and one billion barrels have been thus recovered between 1919 and 1934. Phenol has also been used.

In chemical and toxicological analysis where an emulsion with chloroform has been stabilized by organic protein material, addition of a few drops of alcohol is more effective than addition of electrolytes.

Uses of Emulsions

Emulsions are commonly employed for the sake of economy, but even more important, for efficiency. They enable a dilution of water-immiscible liquids to be accomplished, but this leads also to increased ease of handling and to control of viscosity. A concentrate may be prepared which may be diluted as required for use.

Emulsions present a greatly increased area of contact which is of importance in cosmetics where also penetration of the skin is promoted. In general, they permit uniform spreading, coating, or spraying on surfaces, as in applying paints and adhesives or in the spraying of roads for dust. In foodstuffs digestibility is improved, as in gravies, mayonnaise, or homogenized milk, and emulsions lend themselves to modifying the taste, greasiness, and other properties of liquids. It is possible to emulsify solids above the melting point, cool the emulsion, and keep the material in suspension, as with commercial wax emulsions and with mixes for quick biscuits, pastry, pancakes, etc.

FOAM

Foams are similar to creamed emulsions except that the droplets are replaced by bubbles, and hence gravity has a much greater influence. Further, the stabilizing agent can be stored only in the continuous phase and in the interface, but not in the bubbles. Likewise, the interchangeable terms "foam" and "froth" are restricted to systems in which the gas is the enclosed or discontinuous phase. The terms cease to be applicable when the volume of air is appreciably exceeded by the volume of liquid. If the bubbles are submerged, as in a wet foam, they remain nearly spherical and are not separated by films. Wet foams are formed when the air content is being renewed by bubbling or beating, etc., and when the height of the foam and the fluidity of the liquid are insufficient to effect thorough drainage. If the rate of breakage of the exposed upper surface of the foam is sufficiently great as compared with drainage, the foam will be kept wet until it disappears. A dry foam usually contains less than 10% of liquid by volume. A dynamic foam is one to which new bubbles are being added, and a steady state is reached when the rate of breakage equals the rate of replenishment. A static foam is one that has no mechanism of replenishment as it subsides.

Foaming volume may be defined as the total volume to which one volume of liquid is expanded when it is being turned into foam by a

given method under given conditions. It is small when all bubbles break and escape upon reaching the surface. It is larger when all the liquid may be bubbled into stable foam, and even this can be further expanded by incorporating some larger bubbles in the already formed foam, although soon such bubbles may channel and escape. There is no direct relation between foaming volume and the stability of the foam produced.

Foam stability is measured by study of a dynamic or a static foam. It is expressed for various purposes as: time for total collapse of the foam; average life of the gas in the foam, L_g; average life of the gas in the foam divided by the height of the column of foam in cm, L_g/h; average life of the liquid in the foam, L_l, where L_l is usually quite different from L_g. The drainage of liquid from a foam, on whose rate L_l depends, begins at once, whereas with many foams, including that of oil, the escape of gas may be delayed and then proceed, as is shown in Figure 2·11. Foams may be stable for only a few sec-

FIGURE 2·11 Diagrammatic sketch showing the composition of a froth on standing, indicating the amount of gas in the foam and the amount of liquid in the foam, plotted against time.

onds or minutes or may be so stable that if protected they can be kept for months.

Pure liquids do not foam; the bubbles break as soon as they reach the surface. Mixtures of similar hydrocarbons do not foam, but mixtures of paraffins with aromatic hydrocarbons do foam. Hence lubricating oils foam. The ease with which foam is formed depends not only upon the liquid and the impurities or substances dissolved in it but also upon the means employed for producing the foam, such as beating, spraying, splashing, shaking, of which incomparably the most efficacious is to bubble air in through a finely porous diaphragm or "stone" or sintered glass. A 10% solution of ethylene glycol in water froths when it is bubbled, but the froth is destroyed or does not form on beating.

Once again attention is drawn to the role of a frothing agent, analogous to that of an emulsifying agent or a stabilizing agent in a colloidal sol or the soluble groups that hold an otherwise insoluble film spread on water. In each case the agent is spread in an interface so thinly that its surface properties far outweigh its bulk properties.

Theories

There is no general theory of foams. Foulk's [55] balanced layer theory, however, emphasizes the difference in concentration between the surface and the bulk of the liquid in equilibrium with it. Work is required to eliminate such differences in a film formed by the approach to each other of two existing liquid surfaces, as when two bubbles come together or a bubble rises to a free surface. This may be generalized by pointing out that the spontaneously produced difference between the bulk and the surface zone of a liquid must require work to eliminate it again, and hence will tend to confer foam stability. This generalization furnishes no explicit mechanism but supplements Foulk's idea of film stabilization due to positive or negative adsorption.

Hardy [56] found that an insoluble film on water corresponding to about one half of a complete condensed monolayer affords a maximum in film stability, where the surface pressure (Chap. 3) is only one or two dynes. Similarly other authors have found that a partially saturated monolayer of soluble material adsorbed at the surface produces a good foaming agent. A 0.19 N sodium oleate solution, which is the concentration for the greatest lowering of surface tension, gives the longest-lived foam and below 0.00004 N no foam is obtained. Hardy demonstrated that a film may consist of comparatively immobile monolayers of oleic acid between which lies aqueous liquid that will flow under the influence of an electric field. A. P. Brady points out that the diffuse electrical double layer present can account for this and its stability.

Talmud [57] combines the oriented monolayer ideas and experiments of Hardy with a deep surface arising from hydration of the monolayer. He supposes that the tenacity of an adsorbed film depends upon the degree of hydration of its polar groups, which is assumed to be at a maximum when the monolayer is about half saturated, because here the polar groups are most directly presented to the water without being so closely squeezed that there is no room for the water of hydration. This interlocking, hydrated structure is supposed to resemble a gel.

Talmud also found that drainage is slowest when the hydration is most complete. Further, the stability of a bubble increased if the solvent were evaporating but decreased if the solute were evaporating.

Hardy and McBain suggested independently [58] in 1927 that an adsorbed monolayer must serve as an anchor for chains of oriented molecules extended inward into the liquid like iron filings held by a magnet. Such chains would be transitory, continually forming and

reforming, but would have a distinct tendency to immobilize undisturbed surfaces. This concept should be extended to include polar orientation or increased "cybotaxis" of solvent molecules in a similar manner, likewise tending to rigidify the surface zone. A further factor is the diffuse electric double layer which may extend out to a distance of thousands of Ångström units from the surface. Many observers have reported that the neighborhood of a surface is much less mobile than the bulk of the liquid. Dr. J. V. Robinson, in the author's laboratory, has found that a bubble formed in a lubricating oil may rise with progressive slowness as its surface develops or ages. Chapter 4 is devoted to this subject.

Finally, Wilson and Ries [59] extended the idea of a rigid surface layer, extending beyond monolayers to sorbed layers of great thickness, even up to 40 microns, easily visible in the microscope. They obtained these with aqueous soap solutions, and such layers are familiar in protein foams such as marshmallow whips and meringues. They were able to demonstrate the gradual thickening and jellification of the zone under the visible surface. Foams are often promoted by imperfectly wet solid powders. The previous treatment of the surface of the powders conditions them so that they may promote either foaming or defoaming. Clay promotes the frothing of sodium oleate solutions otherwise too dilute to lather.

In unstable foams the life of the foam is primarily dependent upon the viscosity of the bulk liquid. Thus lubricating oils whose foam lasted for half an hour at room temperature lost all their froth within one minute at 100° C., the life of the foam being proportional to the kinematic viscosity of the oil. Likewise, the life increases rapidly with fineness of dispersion. For purely hydrodynamic movement the rate of rise of a bubble through a liquid should be inversely proportional to the square of its diameter.

Industrial Applications

Most industries have two foam problems. One is to promote desirable foams and the other, just as important, is to eliminate undesirable foams. Desirable foams are familiar in soap systems such as shampoos, etc. In fire fighting a thick blanket of bubbles may coat even vertical surfaces; such foams are stabilized by soybean products or basic ferric or aluminum sulfate and are produced by spraying through a special foam nozzle. Such a nozzle can produce 4500 gallons of foam a minute and throw it 150 feet. Beer and soaps are popularly judged by their

foam, although in the latter case this bears little relation to detergent action. In cooking and baking froths are important.

The enormous flotation industry [60] is based upon control of frothing. For this purpose a moderately unstable foam is produced by a *frothing agent*, and the solid particles to be floated must be imperfectly wetted, which requires conditioning of their surfaces. For example, cresol provides polar groups, and kerosene causes imperfect wetting and makes a highly effective combination with powdered ore. Hence, in flotation a frothing agent is used which, however, tends to lower the surface tension. A *collector* is used to decrease selectively the wetting of the solid, making it more hydrophobic. Pine oil is a collector which only slightly lowers the surface tension. Excess of frothing agent and of collector is avoided. Flotation is further conditioned by the use of *regulators, depressing agents,* and *activating agents.* Often a maximum effect is observed at the isoelectric point.

Permanent foams produced in natural and synthetic rubber, cellulose, gelatin, resins, and glass have great commercial importance. On the other hand, in the glue and paper industries, in the processing of foodstuffs, in the disposal of sewage, or in the dehydration of oil, foams may produce many extremely undesirable complications.

The most effective method of breaking a foam is often to destroy the foaming agent or else to make the film nonuniform by exposure to some defoaming agent or to thermal shock through a blast of cold or hot air or sudden changes of pressure. Acids and other electrolytes may be used to destroy or neutralize the effects of soaps. Chlorination inactivates organic impurities, a famous case being the use of lead strips on steel containers in Deep Springs Valley, California, to form an electric circuit in the brine. Familiar antifoaming agents which may destroy large masses of foam are traces of such substances as ether, isoamyl alcohol, caprylic or nonylic acids, and isoamyl valerate. Capryl alcohol (2-octanol) has recently been specially recommended on account of its insolubility and efficiency and avoidance of the bumping caused by normal octyl alcohol during boiling.

It is well to remember that many of these antifoaming agents are of influence only when they are freshly presented to the foam. Thus foam on beer collapses if there is any ether vapor in the room, but beer foams perfectly freely if ether is first put into the beer and then the foam is produced. Some antifoaming agents are soluble and are specific to particular foams. Frequently antifoaming agents lose their efficacy at higher temperatures. Turkey red oil, however, is an example of an effective soluble antifoaming agent at boiling temperatures.

Many of the most effective antifoaming agents are insoluble but highly emulsified.[61] The emulsion droplets when touching adjacent bubbles cause them to coalesce and also to burst at the surface. Such agents can be extraordinarily effective in minute concentrations. However, they may leave a very small amount of very stable residual foam where they have been spread upon bubbles before these have had an opportunity to burst. Dr. J. V. Robinson and W. W. Woods [62] have shown how insoluble agents may be selected from consideration of their surface and interfacial tensions. Octyl alcohol is a favorite defoaming agent for aqueous systems. Its action is probably as follows: In combination with a detergent it lowers surface tension much more than either alone. The higher alcohols are insoluble, and when a minute droplet is presented or comes into contact with a film or bubble, the surface tension is locally greatly reduced, producing a violent movement of the surrounding film in an effort to equalize the surface tension, resulting in breaking of the bubble.

The author has noticed an intriguing parallel among the substances used for breaking foam, those used for tanning leather, and those used for the treatment of the thixotropic muds used in drilling wells for petroleum.

Systematic compilations of references to the physical chemistry of foams have been made by Sophia Berkman and Gustav Egloff.[63]

REFERENCES

Emulsions and Foam

ALADIN, *Technisch Verwendbare Emulsionen*, Allgemeiner Industrie Verlag, G.m.b.H., Berlin (1928)

ATKIN, W. R., et al. (21 authors), *Emulsion Technology, Theoretical and Applied*, 2nd ed., Chemical Publishing Co., Inc., Brooklyn (1946)

BENNETT, H., *Practical Emulsions*, Chemical Publishing Co., Inc., Brooklyn (1943)

BERKMAN, S. and EGLOFF, G., *Emulsions and Foams*, Reinhold Publishing Corp., New York (1941)

CLARK, N. O., *A Study of Mechanically Produced Foam for Combating Petrol Fires*, Chemistry Research Special Report #6, His Majesty's Stationery Office, London (1947)

CLAYTON, WILLIAM, *The Theory of Emulsions and Their Technical Treatment*, 4th ed., Churchill, London; P. Blakiston's Son & Co., Inc., Philadelphia (1943). Also available in Russian and German

JOHN, W. D., *Modern Polishes and Specialties*, Chemical Publishing Co., Inc., Brooklyn (1947)

SUTHEIM, G. M., *Introduction to Emulsions*, Chemical Publishing Co., Inc., Brooklyn (1946)

Text References

1. CLAYTON, *op. cit.*
2. SUTHEIM, *op. cit.*
3. ATKIN, *op. cit.*
4. HOLMES, H. N., *J. Phys. Chem.*, **29**, 1205 (1925)
5. THOMAS, A. W., *J. Am. Leather Chem. Assoc.*, **22**, 171 (1927)
6. NEVILLE, H. A., *Am. Dyestuff Repr.*, **34**, 534 (1945)
7. A suggestion by Dr. Hans Cassel at the California Section of the American Chemical Society, December, 1935, to explain Sir William Hardy's curious case of an aqueous emulsion where the droplets of oil repeatedly returned to a diameter of two millimeters (*Nature*, 107, 228 (1921))
8. SCHULMAN, J. H. and HOAR, T. P., *Nature*, **152**, 102 (1943)
   SCHULMAN, J. H. and MCROBERTS, T. S., *Trans. Faraday Soc.*, **42B**, 165 (1946)
9. WINSOR, P. A., *Trans. Faraday Soc.*, **44**, 376 (1948)
10. NEWMAN, F. R., *J. Phys. Chem.*, **18**, 34 (1914)
11. CARRIERE, *Chem. Weekblad*, **26**, 413 (1929); VANDER BURG, *ibid.*, **28**, 251 (1931)
12. BRIGGS, T. R., *J. Phys. Chem.*, **24**, 120 (1920); **19**, 478 (1915)
13. TRAVIS, P. M., *Mechanochemistry of the Colloid Mill*, Reinhold Publishing Corp., New York (1931)
14. BONDY, C. and SOLLNER, K., *Trans. Farad. Soc.*, **31**, 835, 843 (1935); **32**, 556 (1936)
15. Striking lecture experiments by CALDWELL, *J. Chem. Ed.*, **12**, 444 (1935). Similarly with throwing alcohol solution of vaseline or acetone solution of fat into water.
16. GAD, J., *Arch. Anat. u. Physiol.*, Leipzig, 181 (1878)
17. QUINCKE, *Wiedeman's Ann.*, **35**, 593 (1888)
18. GURWITSCH, L., *Wissenschaftliche Grundlagen der Erdölebearbeitung*, p. 200 (1913) Translation by Moore, Chapman, and Hall, Ltd., London, p. 430 (1932)
19. RASCHEVSKY, N., *Z. Physik*, **46**, 568 (1927); **48**, 513 (1928)
20. MCBAIN, J. W. and WOO, T. M., *Proc. Roy. Soc.*, **163A**, 182 (1937); see also COFMAN, V., *Chem. Rev.*, **4**, 1, 14 (1927)
21. HART, R., *Ind. Eng. Chem.*, **21**, 87 (1929)
22. KAMINSKI, A., M.A. Thesis, Stanford University, December 1946
23. MCBAIN, J. W. and KAMINSKI, A., *Proc. Roy. Soc.*, **198A**, 447 (1949)
24. MCBAIN and WOO, *loc. cit.*
25. An excellent discussion of the numerical relations of four kinds of close packing of spheres is given by E. MANEGOLD, R. HOFMANN, and K. SOLF, *Kolloid Z.*, **56**, 142 (1931)
26. PICKERING, S. U., *J. Chem. Soc.*, **91**, 2002 (1907). Lawrence has repeatedly emphasized that Pickering's system was not entirely or strictly an emulsion, but this objection does not apply to the experiment as commonly repeated. See E. A. HAUSER and J. E. LYNN, *Experiments in Colloid Chemistry*, McGraw-Hill Book Co., Inc., New York and London, p. 125 (1940)
27. PARKE, J. B., *J. Chem. Soc.*, 1458 (1933); 1112 (1934); WILSON, C. L., *ibid.*, 1360 (1934)
28. WOLFF, H. and ZEIDLER, G., *Kolloid Z.*, **74**, 103 (1936)
29. PICKERING, S. U., *loc. cit.*
30. BENNISTER, H. L., *Soap, Perf. & Cos.*, **13**, 191 (1940)
31. GATTEFOSSE, R. M., *Soap, Perf. & Cos.*, **12**, 961 (1939)
32. SERRALLACH, J. A. and JONES, G., *Ind. Eng. Chem.*, **23**, 1016 (1931); **25**, 816 (1933)
33. HARDY, SIR WILLIAM, *Proc. Roy. Soc.*, **86A**, 634 (1912); **88A**, 330 (1913); *J. Gen. Physiol.*, **8**, 641 (1927); REYCHLER, A., *Kolloid Z.*, **13**, 252 (1913)
34. HOLMES, H. N. and WILLIAMS, H. A., *Coll. Symp. Monog.*, **II**, 135 (1925)
35. HARDY, SIR WILLIAM, *Nature*, **107**, 227 (1921)

36. DEUTSCH, D., *Ber.*, **60**, 1036 (1927); *Z. physikal. Chem.*, **136**, 353 (1928); MEE, A. J., *School Science Rev.*, **17**, 298 (1935-36)

37. WESTON, F. E., *Chem. Age* (London), **4**, 604, 638 (1921)

38. SCHULMAN, J. H., *et al.*, *Proc. Roy. Soc.*, **122B**, 29, 46 (1937); **126B**, 356 (1938); *Trans. Farad. Soc.*, **35**, 716 (1939); **36**, 651 (1940); *Nature*, **144**, 100 (1939); TRIM, A. R. and ALEXANDER, A. E., *Nature*, **154**, 177 (1944); DANIELLI, F., *Nature*, **156**, 468 (1945)

39. DEUTSCH, *loc. cit.*

40. CASSEL, H., *Nature*, **137**, 405 (1936)

41. CHEESEMAN, D. F., KING, A., and SUGDEN, J. N., *Nature*, **141**, 1099 (1938)

42. MCBAIN, J. W. and MCCLATCHIE, W., *J. Am. Chem. Soc.*, **54**, 3266 (1932)

43. TALMUD, D. L. and SUCHOWOLSKAYA, S. D., *Z. physikal. Chem.*, **154**, 277 (1931); TALMUD, B. A., TALMUD, D. L. and KHARITON, YU. B., *J. Phys. Chem.* (USSR), **5**, 107 (1934)

44. LEWIS, W. C. M., *Trans. Farad. Soc.*, **30**, 958 (1934)

45. POWIS, F., *Z. physikal. Chem.*, **89**, 186 (1914)

46. KING, A., *Manufacturing Chemist, Manufacturing Perfumer*, Sept. 1941, p. 213

47. SUTHEIM, *op. cit.*, p. 132

48. REHBINDER, P. A., GOLDENBERG, N. L. and AB, G. A., *Koll. Zhur.*, **9**(1), 67 (1947); through *Chem. Abs.*, **41**, 4352 (1947)

49. WELLMAN, V. E. and TARTAR, H. V., *J. Phys. Chem.*, **34**, 379 (1930)

50. AHERNE, J. C. and REILLY, J., *Nature*, **154**, 86 (1944)

51. MERRILL, R. C., *Ind. Eng. Chem.* (Anal. Ed.), **15**, 743 (1943)

52. KING, A. and MUKHERJEE, L. N., *J. Soc. Chem. Ind.*, **58**, 243 (1939); **59**, 185 (1940)

53. LOTZKAR, H. and MACLAY, W. D., *Ind. Eng. Chem.*, **35**, 1294 (1943)

54. See the chapter on crude petroleum emulsions by R. KOETSCHAN in *Kolloid-chemisches Technologie*, Teil 6, Theodor Steinkopff, Dresden and Leipzig (1932)

55. FOULK, C. W. and MILLER, J. N., *Ind. Eng. Chem.*, **23**, 1283 (1931); FOULK, C. W., *ibid.*, **33**, 1086 (1941); FOULK, C. W. and BARKLEY, J. E., *ibid.*, **35**, 1013 (1943)

56. HARDY, W. B., *Proc. Roy. Soc.*, **86A**, 610 (1912); *J. Chem. Soc.*, **127**, 1207 (1925); TRAPEZNIKOV, A. A., *Acta Physicochim.*, **13**, 265 (1940); BROWN, A. G., Stanford Research Institute, not yet published.

57. TALMUD, D. L. and SUCHOWOLSKAYA, S. D., *loc. cit.*

58. HARDY, W. B., *J. Gen. Physiol.*, **8**, 641 (1927); MCBAIN, J. W. and DAVIES, G. P., *J. Am. Chem. Soc.*, **49**, 2230 (1927)

59. WILSON, R. E. and RIES, E. D., *Colloid Symp. Monog.*, I, 145 (1923)

60. GAUDIN, A. M., *Flotation*, McGraw-Hill Book Co., Inc., New York (1932)
PETERSEN, W., *Schwimmaufbereitung*, Theodor Steinkopff, Dresden and Leipzig (1936)
RICHARD, T. A., *Concentration by Flotation*, John Wiley & Sons, Inc., New York (1921)
WARK, I. W., *Principles of Flotation*, Australasian Institute of Mining & Metallurgy (Inc.), Melbourne (1938)

61. ROSS, S., and MCBAIN, J. W., *Ind. Eng. Chem.*, **36**, 570 (1944)

62. ROBINSON, J. V. and WOODS, W. W., *J. Soc. Chem. Ind.*, **67**, 361 (1948)

63. BERKMAN, S. and EGLOFF, G., *op. cit.*; *Chem. Rev.*, **15**, 377-424 (1934)

CHAPTER *3*

SORPTION

Sorption comprises a number of phenomena,[1] the most commonly recognized of which is accumulation of one or more substances in the interface. However, orientation of the molecules in the interface is of equal importance. Combination of the sorbed substance with a solid can take place in an interface by involving either strong chemical valence bonds or weaker physical interaction through van der Waals' forces. The first is termed "chemisorption," the latter "physical type sorption." The former is confined to surface atoms, but the latter may involve many successive layers of molecules.

Sorption occurs wherever there is a surface or an interface. The importance of the surface has already been emphasized in describing emulsions.

Although sorption occurs at all surfaces, the most important, and also the most complex, are the interfaces between solutions and solids, such as the surfaces of all colloidal particles in any liquid medium; and, therefore, the study of interfaces between liquids and solids is most significant for the study of colloids.

The substances sorbed in such interfaces are sometimes products of reaction or of hydrolysis, sometimes unaltered molecules, and sometimes particular ions. The latter cause electrification of the surface, but surface potentials or differences in potential between contiguous phases may also be conditioned by oriented sorbed dipole molecules either of the surface itself or of the sorbate. Surface tension at the interface may be greatly altered by sorption, and the external properties of a liquid or solid are correspondingly modified, sometimes to a surprising extent.

Sorption is often complicated, not only by diffusion into a solid but also by the property of surface mobility exhibited by many sorbed molecules whereby they are able to move about freely on the surface, including any pores or channels in the solid.

The forces involved in sorption are considerable. When oleic acid on galena is displaced by amyl alcohol, the reaction is so violent that the surface of the galena is disrupted.[2]

Orientation of Molecules in the Surface

To emphasize the importance of the orientation of surface molecules, reference may be made to studies of sorption in which solid palmitic acid is used to sorb gaseous ammonia. H. Devaux [3] had found that when a lens of molten, fatty acid floating on hot water is cooled and solidified, the upper surface is like paraffin and not wettable by water, whereas the lower surface remains freely wettable. He later achieved the same effect with films of silver and copper [4] produced chemically on the surface of water. The Talmuds [5] employed two specimens of finely divided palmitic acid, each with one of these surfaces. The wettable surface with the carboxyl groups outwards sorbed 138 times as much ammonia as the ordinary surface with the paraffin chains outwards and 138,000 times as much as a surface of paraffin wax. Adsorption on the exposed carboxyl groups amounts to a very large number of layers of ammonia.

Another example of orientation in adsorption is afforded by silver bromide or silica which sorb a monomolecular layer of gelatin held irreversibly by its polar groups,[6] leaving the nonpolar groups exposed. These sorb reversibly another layer of gelatin whose now-exposed polar groups confer on the solid many of the properties of protein itself. A similar explanation is given for the sorption of detergent ions on proteins. Completion of the first monolayer brings the solubility to a minimum and restores it as soon as the second monolayer is laid on with the polar groups outwards.[7]

Sorption of Gases and Vapors by Solids [8]

The simplest case for theoretical study is a surface between a gas or vapor and a solid. De Saussure in 1814 showed that all sorts of gases are taken up (sorbed) by all sorts of solids, that the amount of the physical type of sorption is parallel with the condensability of the vapor, and that heat is evolved during sorption.

Küster in 1894 in attempting to apply the mass law to sorption set up instead the empirical classical isotherm

$$x/m = k\,c^{1/n}$$

where x is the weight of substance sorbed by m grams of solid, k and n are empirical constants varying with temperature, n being greater than 1 except at high temperatures. Usually, instead of concentration c, the pressure p of the gas or vapor is employed. In most cases a better agreement is obtained with the Langmuir isotherm

$$x/m = \frac{a\,b\,p}{1 + a\,p}$$

where a and b are empirical constants at any one temperature. This equation, derived for a particular type of plane surface, is found to apply empirically to the sorption of gases and vapors by most solids, including highly porous bodies.

The physical type of sorption is pronounced at low temperatures and disappears at very high temperatures; but chemisorption, while it is most noticeable at moderate temperatures, can sometimes occur at the lowest temperatures as well as at a white heat. Chemisorption is of essential importance in the catalysis of gaseous reactions and is much more specific than physical types of sorption.
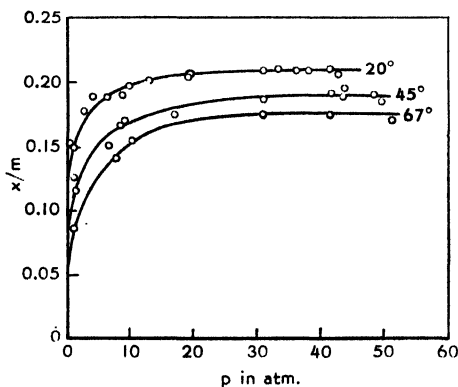
FIGURE 3·1 Typical form of isotherm for all cases of chemisorption and most cases of physical sorption.

Figure 3·1 exhibits a typical form of the isotherm for all cases of chemisorption and for most cases of physical type sorption, such as the sorption of argon or carbon dioxide by charcoal at a given temperature. With rise of temperature, keeping pressure constant (isobar),

sorption usually diminishes steadily. However, sometimes the isobar resembles that diagrammatically shown in Figure 3·2. The interpretation given by the Princeton school is that the portion of the curve $A B$ at low temperatures represents physical type sorption, falling off with rise of temperature. Over the range $B C$ appears chemisorption,
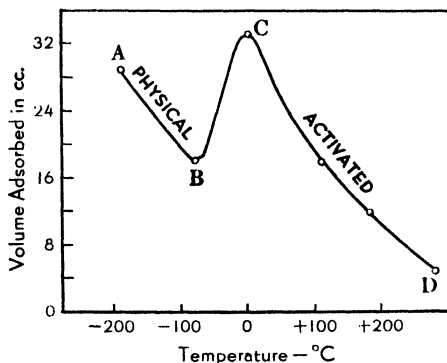
FIGURE 3·2 Isobar of Princeton school. Taken from Taylor and McKinney, J. Am. Chem. Soc., 53, 3604 (1931).

which is supposed to be a slow reaction having a high temperature coefficient. This increases the total sorption, and at the highest temperature range $C D$ this again falls off. However, the work of J. K. Roberts of the Cambridge school in England has shown that in some cases, such as hydrogen with nickel, the hydrogen is instantaneously chemisorbed as single atoms at all temperatures from that of liquid air up. It is therefore chemisorption but apparently not "activated" adsorption. The portion of the curve $B C$ is that ascribed to solution of the gas in the metal, which is slow and has a positive temperature coefficient. The hydrogen atoms move freely on the nickel surface.

When describing sorption of vapors it is better to plot relative pressure, p/p_s, as in Figure 3·3, rather than absolute pressure p,[9] where p_s is the saturated vapor pressure of free liquid at the same temperature. Then the curves for various temperatures look nearly alike except that the amount sorbed at higher temperatures is progressively less in about the same proportion as would be the weight of the corresponding liquid if warmed in a filled but open container of constant volume. This gives the temperature coefficient of such sorption. On clean surfaces of charcoal at almost zero pressure x/m is between a third and a half of the full value for saturation pressure. At saturation pressure, $p/p_s = 1$,
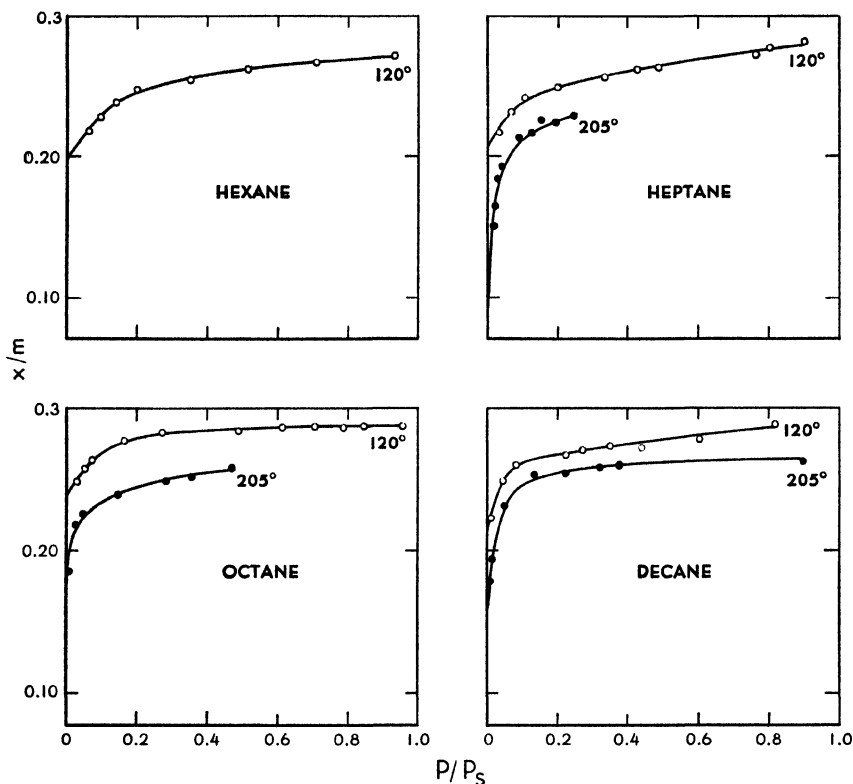
FIGURE 3·3 Adsorption isotherms for vapors by highly evacuated sugar charcoal. Taken from Weiser, H. B., *Colloid Chemistry*, p. 30, John Wiley and Sons, New York (1939).

roughly equal volumes of various liquids are taken up by the same sample of charcoal.

Many attempts have been made to determine the effective surface area of porous bodies. The procedure mostly in current use is that introduced by Brunauer, Emmett, and Teller, generally referred to as the B.E.T.[10] method. An adsorption isotherm similar to that of Figure 3·4 is obtained for physically sorbed nitrogen near the boiling point of liquid nitrogen. At point B it is assumed that the accessible surface is covered with a monolayer of nitrogen molecules and from this is calculated the area. Areas so calculated are in general agreement with those obtained by some other methods, such as the absolute method of Harkins and Jura, who measured heat of emersion of a saturated nonporous powder.[11] Above the pressure at B, multilayer sorption is supposed to occur. It is held preferable in some cases to make use of the B.E.T. straight line equation.[12] Harkins and Jura have derived a

simple method of plotting $\log p/p_s$ against the reciprocal of the square of the volume adsorbed $(1/V^2)$. From the slope, which they have standardized for the use of nitrogen, water, n-butane, and n-heptane, the area is determined. Ross [13] finds that Hüttig's alternative to the B.E.T. equation describes the experimental results more closely.

Dole [14] points out that the B.E.T. equation gives the number of adsorption sites, not the surface, which latter can only be obtained by a further calculation in which a value for the area per site is assumed. He then describes the conditions under which multilayer sorption can lead to various modifications of the B.E.T. equation.[15]

Attempts to measure surface areas by sorption from solution usually yield an underestimate due to preemption of part of the surface by solvent and sometimes also to inaccessibility of very fine pores to large molecules.[16] However, the negative adsorption of the ions of electrolytes has been used to measure surface areas, a method which possesses the advantage of not having to alter the solid by exposing it to high vacuum as in the B.E.T. procedure.[17]
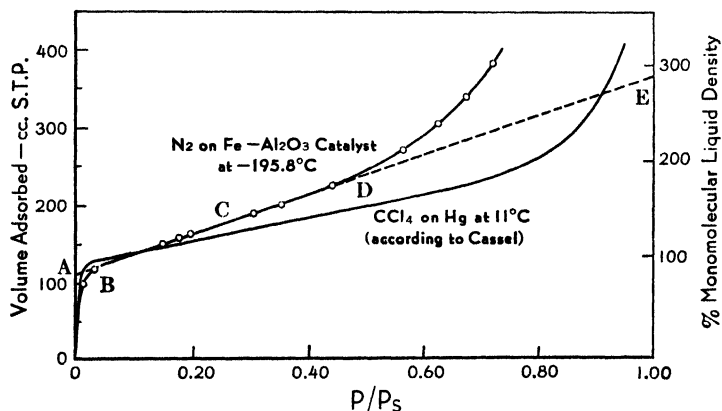
FIGURE 3-4 Comparison of sorption isotherm for nitrogen at $-195.8°$ obtained on iron catalyst, and (calculated) for carbon tetrachloride on liquid mercury. Taken from Emmett and Brunauer, *J. Am. Chem. Soc.*, **59**, 1558 (1937).

The areas involved in a porous body are enormous; for example, 2.6 million square meters in 1 kilogram of charcoal. Many properties, such as the hardness and tensile strength of metals or the bending strength of glass, are not entirely bulk properties but are often conditioned by sorption of surface active materials or even of air.

Rehbinder [18] has shown that the hardness of many solids can be altered by sorption upon them; this he has used to increase the speed of drilling and boring operations. Engelhardt [19] mentions for example

that soap solutions which were effective with metals actually impeded the grinding of porcelain, whereas a more suitably chosen reagent greatly accelerated it.

Sorption of gases by solids has been put to practical use in many directions: the production of high vacua, the liquefaction of gases, the separation of gases from gaseous mixtures, the recovery of vapors as in rayon manufacture, and even the isolation of para-hydrogen from liquid hydrogen.

Sorption from Solution

Sorption from solution is readily demonstrated by exposing a dye solution to a solid and noting how the solution is robbed of the dye, the solid becoming deeply colored in the process. If known quantities of solid and solution are taken and the final concentration of the solution is compared with the original, the loss from the solution so calculated is taken as a measure of the amount of sorption by the solid. It will be noted that this assumes that the solvent itself is not sorbed at all, whereas in fact solvent and solute compete for the available surface.

Another striking demonstration of sorption is that suggested by Tanner [20] in which a large clean glass flask is exposed as briefly as possible to a solution of malachite green and then rinsed many times with large portions of tap water until the rinsings appear completely colorless. Then the flask is rinsed once with a small amount of ethyl alcohol, which becomes bright green. This rinsing is similarly effective if done with a soap solution or with a concentrated solution of salt. This result illustrates the fact that sorption of dye by the glass begins quickly and that the degree of sorption is different in different solvents, and, further, that ion exchange may enter into sorption. In fact, the dye base sorbed is not identical with the dye salt to which the glass was originally exposed, preferential sorption and hydrolysis having taken place.

A series of beautifully simple experiments to demonstrate sorption in a liquid-liquid interface has been described by Deutsch.[21] Solutions of various indicators or dyes in benzene, resting upon a layer of water, are shaken, whereupon the temporary emulsion assumes a new color, the color disappearing again as soon as the interface is destroyed by separation of the water and benzene. This may be repeated again and again. Similarly, if powdered quartz and filter paper are exposed to a colorless benzene solution of rhodamine they are colored deep red.

Foam on a yellow solution of thymol sulfophthalein is colored red, but after collapse the liquid is yellow again.

Sorption from solution is much more universally observed than sorption of gases, and it is the essential phenomenon in many industries, such as dyeing, tanning, the complex industrial treatment of textiles, detergency, decolorizing, defecation. It is essential in the preparation and separation of hormones, vitamins, and enzymes. It governs the properties of such materials as clay. It is responsible for the blue color of starch with iodine and of similar blues with other solids, such as lanthanum hydroxide or barium sulfate. It affects crystallization and positive and negative catalysts. It is responsible for the chief properties of colloids. It is greatly influenced by pH and by whether the solid is positively or negatively charged in contact with the solution. It is often highly specific, but in the paragraph on chromatography it will be shown how the slightest difference in sorption can be magnified to produce quantitative separation of similar substances. Sorption from solutions, in contradistinction to sorption of gases and vapors from the gas phase, is not very greatly affected by temperature or even extreme pressure.

The Classical Empirical Isotherm

Almost universally, for any given temperature, the empirical classical isotherm

$$x/m = k\, c^{1/n}$$

applies with fair accuracy to sorption from solutions over the widest possible concentration range. Here c is the concentration of the solution in *equilibrium* with the solid and does not refer to the original concentration of the solution used. The empirical exponent $1/n$ may have any value between about 0.5 and 0.1. Likewise, the values of the empirical constant k reflect all degrees of sorption from the least up to one that is so nearly complete that the equilibrium concentration is infinitesimal, where rinsing removes a correspondingly negligible amount of the sorbed material. It is a property of this classical isotherm, especially when $1/n$ has a very low value, that sorption is already pronounced in extreme dilution. It has been shown that the concentration c is a true equilibrium concentration which is reached by both desorption and adsorption for any particular value of x/m. Occasionally, for some unexplained reason, the Langmuir isotherm holds empirically remarkably well for a complex case, as for example the sorption of protein by nitrocellulose.[22]

Effect of Time

The foregoing holds true only for experiments carried out with a definite total time of exposure because usually, owing to the inaccessibility of the fine pores of solids and the competition between solvent and solute, sorption may require long periods, even up to years, for completion. Burt's empirical formulation of this is

$$x/m = k \, t^{1/n}$$

where t is now time, and k and n are empirical constants; but for very prolonged sorption Burt had to modify this equation by including a saturation value σ. The equation now is written

$$\log \frac{\sigma}{\sigma - s} = k \, t^{1/n}$$

where s is the value of x/m for any given time. It is a common experience that a stain may be removed if it is washed off immediately, but delay makes it permanent.

Effect of Nature of Solid, Solvent, and Solute

In general, solids can usually be arranged in a definite order according to their sorbing power. However, there are certain specificities or selectivities. For example, charcoal has much more affinity for hydrocarbons than for water, whereas silica exhibits the reverse behavior. Nerve substance generally resembles silica with respect to sorption.

Comparing solvents, charcoal sorbs organic substances most strongly from water, much less from benzene, and still less from ether and acetone. Sorption is greatest from poor solvents and from solvents with low heat of wetting. The competition for a solid surface between solvent and solute favors the solvent that has the greatest affinity for the solid.

When solutes sorbed by charcoal are compared, organic substances are much more strongly sorbed by charcoal than are inorganic, and aromatic compounds more than aliphatic. Polar organic compounds are strongly sorbed, acids and bases being especially strongly sorbed. However, presence of hydroxyl or amino groups lessens the sorption of organic compounds. This is well illustrated in the refining of sugar with its many hydroxyl groups, in which process the impurities are removed by charcoal, leaving the sugar in solution. Sugar is negatively sorbed even by froth, and this fact has been used in purification of sugar from surface active impurities. A few illustrative measurements are given in the following table: [23]

TABLE 3

SORPTION OF ORGANIC ACIDS AND BASES. EFFECT OF ARRANGEMENT OF
POLAR AND NONPOLAR GROUPS ON ADSORPTION

(cc 0.01 M acid or base sorbed by 0.25 gm blood charcoal from 100 cc of solution)

Acetic	15.68	Benzoic	73.00
Aminoacetic	0.00	o-Hydroxybenzoic	71.78
Chloracetic	27.45	m-Hydroxybenzoic	61.53
Dichloracetic	25.54	p-Hydroxybenzoic	65.36
Trichloracetic	17.78	o-Aminobenzoic	68.53
Butyric	35.39	m-Aminobenzoic	57.32
iso-Butyric	27.42	p-Aminobenzoic	65.15
Valeric	49.39	Sulfosalicylic	56.10
iso-Valeric	44.12	Ammonium hydroxide	0.00
Hexoic	62.13	Tetramethylammonium hydroxide	1.20
iso-Hexoic	59.67	Tetraethylammonium hydroxide	10.07

Strong electrolytes are generally very slightly sorbed and may even appear to be negatively sorbed, except in the cases to be discussed later under the heading "Ion Exchange." Ions raise the surface tension of water in the order $SO_4 > H_2PO_4 > Ac > Cl > NO_3 > Br > I > CNS$, the thiocyanate having scarcely any effect. This is the first example we have seen of the famous Hofmeister or lyotropic series, which is the subject of Chapter 9. It will be noted that in the case of the surface tension of strong electrolytes in water, no colloids are involved.

To avoid too broad a generalization it should be pointed out [24] that at the oil-water interface, acids, amines, and polar molecules capable of ionizing are one thousand times more adsorbable than alcohols, esters, ketones, phenols, and other polar molecules not capable of ionizing at that interface.

With silica in nonaqueous solutions, dinuclear or polynuclear aromatic compounds are strongly sorbed. Next in order come molecules containing a phenyl group. Least strongly sorbed are cycloparaffin and paraffin. Increasing the number of carbon atoms in an n-alkyl group in any molecule always decreases its sorption,[25] as is explained in the next paragraph. Hirschler and Amon [26] have shown the wide applicability of sorption for the separation and purification of such closely related compounds as the saturated hydrocarbons using charcoal and silica. Geometric as well as structural isomers of hydrocarbons have been separated.

Traube's Rule

The essential content of Traube's rule is that in any homologous series molecular weight plays a determining part. In the form given

it by Traube, for fatty acids sorbed by charcoal from water, it was pointed out that to obtain equal amounts of sorption with successive higher members of the homologous series, the requisite concentration is only a half or a third as great as for the next lower homolog. A similar rule holds for the effect on the lowering of the surface tension of water, the highest homologs lowering it most. It is not a contradiction but rather a confirmation of this that Holmes and McKelvey [27] found the order of the homologous series was reversed when charcoal was replaced by silica and water by toluene; the longer the hydrocarbon chain, the stronger the tendency for it to remain in the toluene. A similar result was obtained by Claesson and Claesson [28] in sorption by carbon from hydrocarbon solutions of polymers of different molecular weight, such as nitrocellulose, rubbers, and polyvinyl acetate. Dubinin found apparent exceptions to Traube's rule which were traced to the purely mechanical cause that the pores of certain charcoals were so ultra fine that they were readily accessible only to the smaller molecules. Kiselev, Vorms, and Kiseleva [29] found that the same influence is partly responsible for Holmes and McKelvey's and their own results with silica gel.

It is unexpected that Hutchinson [30] has found no effect whatever of position in the homologous series of alcohols and fatty acids when studying interfacial tension between water and benzene.

Extreme Concentration

Competition between solute and solvent for the surface of the solid becomes apparent in sufficiently high concentration where the ratio of solvent to solute sorbed exceeds their ratio in the original solution, so that the solution becomes still more concentrated and the solute appears to be "negatively" sorbed. It is easily shown that in very dilute solution, on the other hand, sorption of solvent has only a negligible effect upon the measurement of sorption of the solute. Sorption of both solvent and solute is almost never directly determined, but this was done by King and Bakr for iodine and benzene with charcoal.[31]

Sorption from Mixed Solutions

Here the competition is more complicated and the attainment of equilibrium is slower. Naturally, each solute lessens the sorption of the other, but in practice it has been found that a more strongly sorbed solute largely displaces a more weakly sorbed substance from the surface.

Preferential Sorption and Hydrolysis

The matter sorbed upon a surface is often not identical with the substance in the solution presented to it. Frequently an acid or a base is sorbed from a salt solution, hydrolysis having been promoted in order to achieve this. There are many examples of this, such as the sorption of base from basic dyes, setting free hydrochloric acid; the sorption of barium hydroxide from barium chloride solution by starch; the hydrolytic sorption from solutions of compounds of gold; and the sorption of acid soap in foam, or on many solids, setting free a corresponding amount of base. In contrast to its effect on basic dyes, charcoal scarcely sorbs acid dyes.

Interest in this subject received a psychological setback with the demonstration [32] that careful purification of cellulose deprived charcoal of such sorptive power, but interest revived when Bartell and Miller [33] proved that the most highly purified charcoal still can exhibit pronounced, though selective, sorption.

Sorption Conditioned by the Exact Surface of the Solid

The same solid can have totally different sorptive properties when given different pretreatments. It is not merely that the amount of sorption can be enormously increased by "activation," but its selective behavior can be modified or completely reversed.

Activation of charcoal consists in heating in air, steam, or carbon dioxide, with or without previous impregnation with materials such as zinc salts, to burn away more than half of the carbon and leave an ultraporous, skeletonized structure so extreme that practically all the carbon atoms are exposed for sorption. It is as a reminder of this extreme permeation of a porous solid that the author introduced the term "persorption." Some crystal lattices are porous, as in the case of zeolites and ultramarine, so that the term fittingly applies to them also.

The pretreatment of the surface of charcoal is all important. Unactivated wood charcoal sorbs equivalent amounts of anion and cation. Animal charcoal sorbs twice as much anion as an equivalent amount of wood charcoal, but scarcely any cation, but after activation five times more anion is sorbed and still less cation. Conditions of activation therefore determine the sorption observed, as is further demonstrated by the following table, which summarizes the work of a few investigators.

TABLE 4

EFFECT UPON SORPTION OF CONDITIONS OF ACTIVATION OF CHARCOAL

Activation at C°	Investigators	Does It Sorb?	
		Acids	Bases
Air at 1000°	Bartell and Miller [1]	Yes	No
Air at 800°	Dubinin [2]	No	Yes
Air at 550°	Dubinin	Yes	No
O_2 at 1000°	Frumkin [3]	Yes	No
O_2 at 400°	Kruyt [4]	Yes	Yes
O_2 at 1000° then 400°	Kolthoff [5]	No	Yes
CO_2 at 1000°	Frumkin [3]	No	No
CO_2 at 1000° in vacuo	Frumkin	No	No
H_2 at 1000°	Frumkin	No	Yes

1 Bartell and Miller, *J. Am. Chem. Soc.*, **44**, 1866 (1922); **45**, 1106 (1923); Miller, *ibid.*, **46**, 1150 (1924); **47**, 1270 (1925)
2 Dubinin, *Z. physikal. Chem.*, **A140**, 81 (1929); **A150**, 145 (1930)
3 Frumkin and Donde, *Ber.*, **60**, 1816 (1927); Bruns and Frumkin, *Z. physikal. Chem.*, **A141**, 141 (1929); Burstein and Frumkin, *ibid.*, **A141**, 158, 219 (1929)
4 Kruyt, *Koll. Z.*, **47**, 44 (1929); *Kollchem. Beih.*, **32**, 249 (1931)
5 Kolthoff, *J. Am. Chem. Soc.*, **54**, 4473 (1932)

Bartell and Miller's activated charcoal sorbs the acid from a strong salt such as sodium chloride, leaving the inorganic base free. The hydrochloric acid is very tenaciously held. Salts of organic acids are sorbed partly directly and partly by hydrolysis. The organic base tetraethylammonium hydroxide is strongly sorbed. The charge carried by the charcoal in water depends upon the temperature of activation and the atmosphere to which it has been subjected.

Oxygen is chemisorbed by charcoal at room and higher temperatures to form complexes such as $(C_2O_3)_n$ or $(CO)_m$ on the surface.[34] The former sorbs basic organic matter and the latter sorbs iodine.[35]

Charcoal activated at high temperatures is electrokinetically positive in the presence of air or oxygen.[36]

From the foregoing, it is immediately clear that it is not sufficient to give only the name of a solid or of the interior of a colloidal particle, but that the sorptive behavior is conditioned by the actual composition of the *surface*. Furthermore, the ash content of charcoals varies widely in nature and amount and is another major factor in the sorptive behavior.[37]

Ion Exchange

Thomas Way [38] in 1850 found that soils have the power of separating bases from salts dissolved in water. Subsequent work traced this to base exchange between ions in solution and sorbed ions of like sign.

Such ion exchange is the basis of modern soil chemistry.[39] For example, ion exchange is an important part of the measurement of nutrient intake by the roots of plants. Ion exchange has been an essential tool in nuclear chemistry, as for example in the separation of the rare earth elements and many radioactive isotopes.

In base exchange equivalent replaces equivalent. In all cases the classical sorption isotherm governs the equilibria. Heavy metal ions and potassium are preferentially sorbed and require a high concentration of sodium ion to replace them. Base exchange is exhibited by naturally occurring zeolites and montmorillonite clays. It is common in all cases of sorption of electrolytes, even for example with cellulose [40] or glass.[41] Filter paper does not appreciably sorb anything from aqueous solutions except acids, bases, or salts. These may be removed by ion exchange, best of all by surface-active ions, such as those of detergents.

Horowitz and Paneth discovered the general principle that adsorption of an ion is great when the analogous compound of that ion with the element present in the solid is difficultly soluble in the solvent employed.[42]

Phenol, aniline, cresol, or various amines may be condensed with aldehydes to form a highly porous resin whose pores may be lined with hydroxyl, carboxyl, or sulfonic acid groups to form base exchangers, or with amino or other groups [43] to produce anion exchangers or sorbents. Some cheaper ion exchangers are made from sulfonated lignite and other carbonaceous materials.

A base exchanger sorbs cations on the hydroxyl groups and liberates other cations, or an equivalent amount of hydrogen ion, producing a strongly acid solution. If this solution is now exposed to an acid exchanger, or sorber, the anions are removed, leaving no ions in solution except those of water. The anions are held by the amine groups of the second resin. This is a commercial method of producing pure water. It is also an effective method of purifying many colloidal solutions.

Ion exchangers may be regenerated by treatment with cheap acids and bases, respectively, or, if mere base exchange is required, as in domestic water softeners, common salt is used for regeneration.

Other important commercial applications [44] are the purification of molasses and sugar, protein solutions, biological and pharmaceutical media, and the recovery of traces or the removal of objectionable traces from various solutions.

Ion exchange is the basis of the dyeing of certain textiles by anion dyes, such as Congo red, a sodium salt, and by cation dyes, such as malachite green, a chloride. It is also used in the removal of formic

acid from formaldehyde, the artificial ageing of whisky, the recovery of metal from dilute solution, the production of streptomycin, and the isolation of amino products. Ion exchangers are also used in the recovery of alkaloids, the removal of sulfuric acid from sulfonated oils, the reduction of calcium content of milk for infant feeding, the recovery of pectin from grapefruit peels, the removal of acid from dextrose, the removal of sulfuric acid from ethylene glycol, and the processing of sorbitol.

In summary it should be pointed out that ion exchange is not the only way in which ions are sorbed. They may be sorbed by oppositely charged ions already on the surface, whether there by previous sorption or as ions which are part of the crystal lattice of the solid itself. Again, true sorption of the salt itself may occur.

For sorption of ions on crystal lattices Fajans has formulated the following rules based upon Paneth's [45] observation of the strong sorption of an ion if its compound with the oppositely charged ion of the lattice is insoluble. An ion is strongly adsorbed on an equivalent compound of the salt type only when it forms a difficultly soluble or weakly ionized compound with the oppositely charged ion of the lattice. The adsorption of a cation is raised by adsorbed anions, that is, by charging the surface negatively, and is lowered by adsorbed cations, that is, by charging the surface positively. The adsorption of anions is raised by adsorbed cations and is lowered by adsorbed anions.

Separation of Sorbed Substances by Chromatography

Chromatography is a method invented by Tswett in 1906 for enhancing even the minutest difference in sorbability of different substances. It has become so important that a series of books [46] has appeared on the subject, and a remarkable variety of modifications and refinements has been developed. It is practiced as well on a micro scale as with the use of large industrial towers.

The standard method utilizes a carefully packed column of a white solid, such as magnesium oxide, aluminum oxide, sucrose, starch, cotton, 8-hydroxy quinoline, etc. The solvent used in the solution poured into the column should be one from which sorption occurs, but not too strongly. For example, sorption is greatest from petroleum ether and least from pyridine, acids, bases, and alcohol. The column is "developed" by pouring in more solvent, or a more polar solvent, or an "eluant," which separates the less strongly sorbed materials by pushing them progressively down the column. When they are com-

pletely separated, further eluant may wash out each in turn, or the column can be separated with a knife and the layers can be analyzed. This gives quantitative separation and concentration without loss.

The position of ions or colorless materials in the column can be identified by reagents, for example, hydrogen sulfide; or by fluorescence or by refractive index or by reactions of successive portions of the extruded eluant. Chromatography provides a critical test for homogeneity of chemicals and facilitates the comparison of similar substances, their purification or concentration, the recognition of quantitative separation in complex mixtures, and the determination of molecular structure. It may be used for separating and detecting even such inorganic ions as sodium, potassium, and ammonium.

J. Norton Wilson [47] and E. Glueckauf [48] have presented a detailed theory of chromatography.

Filter paper "partition chromatography" is a somewhat similar and useful tool for the separation of amino acids and for biological studies with radioisotopes.[49] However, its main features are thought to depend not upon sorption but upon partition between two immiscible solvents. A drop of the material is placed on a large sheet of filter paper and water-immiscible solvents are flowed over it in one or more directions. An improved technique is described by Williams and Kirby.[50] As many as 32 amino acids have been simultaneously separated into different positions on the filter paper. Similarly, ordinary filter paper chromatography can separate and identify 40 sugars.[51] Metals such as gold, platinum, palladium, copper, and silver present to the extent of less than 1 mg can be separated quantitatively.[52]

It is only fair to point out that many of these and many other separations by filter paper were anticipated by Goppelsroeder's [53] extensive experiments (1861–1910), following Schönbein. Thus he could detect 0.000000025 mg of methylene blue in 1 cc of solution.

Surface Tension and Sorption

Surface tension is a manifestation of the attraction between neighboring molecules. It is expressed in dynes/cm and is numerically identical with the free surface energy in ergs/cm^2. It is measured by a dozen or more different methods, each of which has been subject to a great deal of criticism. Among the best known are those which measure the extent to which a liquid rises in a perfectly wetted capillary tube;[54] the maximum force required to detach a ring from a surface;[55] the pull of a plate partially immersed at right angles to the surface of a liquid;[56]

the shape of a bubble immersed below a solid surface,[57] of a drop resting upon a nonwetted surface,[58] of an interface pressed by hydrostatic pressure into a heavier liquid,[59] or of the contour of a hanging drop; [60] the pressure required to flatten a bubble or a surface at the end of an open tube; [61] the pressure required to produce bubbles from a tube immersed in a liquid; [62] the weight of drops falling from an orifice of known dimensions.[63] It will be noted that these differ in the extent to which any change in surface tension with time may be studied.

A truly static method is one in which the surface is allowed to age until no further change tends to occur. At the opposite extremes are the attempts at measurement of dynamic or instantaneous surface tension, such as studies of the contours of a freshly formed liquid surface as it falls from an orifice.[64] Few methods are either perfectly static or dynamic, but most occupy an indeterminate position in between.

For pure liquids it is a matter of controversy as to whether dynamic surface tension is experimentally distinguishable from the final static surface tension. Buchwald and König [65] obtained positive differences for a number of liquids, such as water and alcohol, but not for benzene, the dynamic being greater than the static surface tension.

In contrast, for solutions the dynamic surface tension is almost always very different from the static surface tension. The dynamic surface tension corresponds to an instantaneous exposure of the interior of the solution before there has been time for diffusion, sorption, or any rearrangements to occur. For dilute solutions the dynamic surface tension is nearly equal to the surface tension of the pure solvent. Thus, Lord Rayleigh in 1890 [66] found that the dynamic surface tension of a 0.25% solution of sodium oleate was 77 dynes, whereas the static surface tension was only 25 dynes. If only simple diffusion into the surface were involved, just a small fraction of a second would be required for its completion. However, it is frequently observed that long periods of time, up to days or weeks, are required to establish final static surface tension.

It might be mentioned here that if a freshly made surface of a colloidal sol possessing some structure is instantaneously stretched, an "ultradynamic" surface tension may be observed. For example, for a solution of rubber in benzene, the ultradynamic surface tension may be 150 dynes/cm, whereas both dynamic and static surface tension are only 29 dynes.[67] This must be related to the elasticity and other flow properties of the colloid. Similarly, for an aluminum soap in a hydrocarbon the ultradynamic surface tension may be between 26

and 300 dynes when the dynamic and static surface tensions are between 22 and 25 dynes.

The Adsorption Theorem of Willard Gibbs (1875), J. J. Thomson (1888), and E. Warburg (1890)

This theorem relating surface tension and adsorption is often quoted erroneously. It is supposed to state that if the surface or interfacial tension is lowered, the solute will be positively adsorbed; whereas if it is raised, the solute will be negatively sorbed or kept away from the surface. However, in actuality this theorem states that if the surface tension of any binary solution is plotted against the logarithm of the activity of the solute, adsorption will depend solely upon the slope at any given point, and for any given slope it is the same for all solutes regardless of the magnitude of the surface tension. If at a given activity or concentration the surface tension curve happens to be horizontal, there will be no adsorption — that is, no difference whatsoever in composition of the surface layers as compared with the bulk of the solution, no matter whether the surface tension is equal to that of the solvent, is greater than, or lies far below that of the solvent.

The Gibbs theorem formulates the excess or deficiency of each and every component in the surface zone, but these quantities, although strictly interrelated, are not given definite values until some arbitrary mathematical surface of reference is chosen in the surface zone. An infinite number of such choices is possible, but the most common is to define an imaginary one in which the concentration of the available solvent could have been maintained constant on each side; that is, equal to that in the bulk concentration of the liquid on the one side, and that in the bulk vapor on the other. It so happens that this corresponds to the empirical results obtained by the McBain-Bakr sorption balance (without correction for buoyancy) or to sorption from solutions as usually measured where the result carries the implicit assumption that the solvent has not been sorbed.

If as a standard of reference the adsorption of the solvent is thus set equal to zero, the theorem arrived at independently by Thomson (1888) and Gibbs (equation 508) (1878), can be expressed in the form

$$\Gamma = -\frac{d\sigma}{d\mu} = -\frac{d\sigma}{d\overline{F}} = -\frac{1}{RT} \cdot \frac{d\sigma}{d\ln a}$$

or approximately,

$$\Gamma = -\frac{c}{RT} \cdot \frac{d\sigma}{dc}$$

where Γ is the excess or deficit of the solute in the surface zone, σ the surface tension, μ the chemical potential, $\overline{F}$ the partial molal free energy, a the activity, and, most simply, c the concentration. Less often quoted is the simple but rigorous equation 585 of Gibbs which expressed Γ in terms of D, the density of the vapor of the solute above the solution, and p, the vapor pressure of the solute above the solution, so that

$$\Gamma = - D \frac{d\sigma}{dp}$$

The adsorption theorem cannot in practice be applied rigorously to systems involving organic substances because of their insufficient purity and their chemical reactivity in surfaces. However, in simple cases it is found to apply as a first approximation.[68] It does not apply to fresh or moving surfaces.

Actual measurements of adsorption upon liquids have rarely been carried out; most authors are content to calculate them from the adsorption equation without experimental verification.
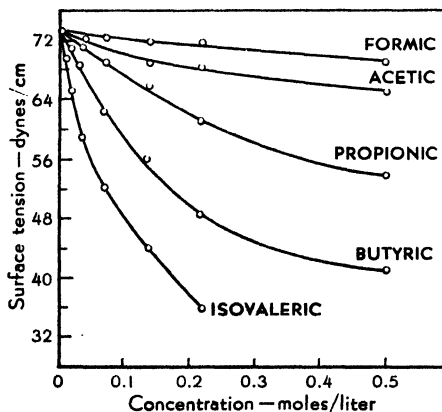
FIGURE 3·5 Lowering of surface tension of water.

Figure 3·5 illustrates the lowering of the surface tension of water by the lower members of the homologous series of fatty acids. It should be noted that the lowering of surface tension is parallel to the increase in molecular weight of a fatty acid within the homologous series. This is in accordance with Traube's rule, which was discussed in connection with sorption.

Surface tension curves may be classified according to the three types [69] shown in Figure 3·6. In Type I surface tension is increasingly

lowered with increasing concentration of solute. This is the commonest case.

In Type II the surface tension of the solvent is increased, although usually not to a very great extent, with increasing concentration of solute. This is observed for many strong electrolytes and for some compounds, such as sugar, containing many hydroxyl groups. It connotes negative adsorption. The outermost layer of a salt solution is pure water.
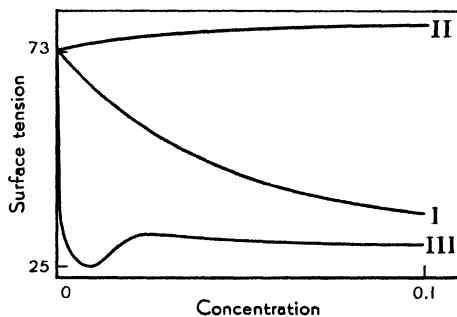
FIGURE 3-6 Classification of surface tension curves.

Type III curves are by no means infrequent, since most soap and detergent solutions exhibit this behavior. It is characterized by a very great lowering of surface tension in very dilute solution, the surface tension thereafter remaining approximately constant, or even passing through a minimum followed by a shallow maximum but still at a surface tension far below that of the solvent. If these were pure two-component systems of water and solute, the adsorption theorem would require that there could be no adsorption at the concentration corresponding to the horizontal portions of the surface tension curve, either at a minimum, maximum, or constant value — nothing would explain the lowering of surface tension. This problem has not been wholly resolved but it has been shown by Miles and Shedlovsky [70] that the minimum is often dependent upon the presence of some other organic impurity, such as a higher alcohol. In the particular instance of the most carefully studied substance, dodecyl sulfonic acid, A. P. Brady [71] has found that the impurity may be temporarily frothed off at the minimum concentration, leaving a purified dodecyl sulfonic acid whose solutions at first show no minimum and whose adsorption is approximately that required by the adsorption theorem.

In contrast to surface tension, the interfacial tension between two liquids is often reduced practically to zero by the addition of a surface

active material. Thereupon, the two liquids can be very readily emulsified. Sometimes, however, the behavior may be similar to that of the three types mentioned.

The adsorption theorem is applicable only to final equilibrium. Direct measurements sometimes fail to find any adsorption after many minutes, although the surface tension has been lowered and is well on its way toward a static value, when the adsorption becomes of the expected order of magnitude.

Frequently the surface tension of a solution is quickly reversible in that it is almost independent of expansion and contraction of the surface.[72] More often, as has been shown in a series of papers from the author's and from Doss' laboratory,[73] the solute returns with some difficulty into solution when the surface is contracted. This has been termed formation of a "pellicle" of ordinarily perfectly soluble material on the surface, which dissolves under the influence of a definite surface pressure. Sometimes, however, denaturation is complete, as with egg albumin, and to a lesser extent with saponin, the solute in the surface remaining insoluble. A bubble or droplet containing saponin in its surface merely wrinkles or shrivels when the air or water in the interior is withdrawn.

There seems to be a potential barrier to be overcome before a solute can take its final place in the surface and produce, after a lapse of time, a final adsorption and static surface tension.

The Surface Tension of Sols

The surface tension of dilute irreversible or unstable sols is usually similar to that of the solvent, as was first pointed out by Linder and Picton in 1905.

It is quite otherwise with the majority of stable colloidal sols. In most cases these show a considerable lowering of surface tension of the sol as compared with pure water. For example, 10% tannin lowers the surface tension of water by 29%, 20% of gum Arabic by about 9%, dilute gelatin and dilute agar by 12.5%, white of egg by 28%, and soap and most detergents by as much as 65%, frequently exhibiting a Type III curve with or without a minimum.

However, silicic acid sol, which has approximately the same surface tension as water, and starch and some gums increase the surface tension of water. Some sodium silicates slightly raise and others slightly lower the surface tension of water.

The interfacial tensions of stable sols against organic liquids some-

times resemble the surface tension curves but are often quite different. For example, with soap solutions the interfacial tension is steadily and asymptotically reduced to zero. When the interfacial tension falls to one dyne or less, the slightest mixing results in emulsification.

Insoluble Films on Water [74]

This subject was pioneered by Lord Rayleigh and Sir William Hardy in England, Miss Agnes Pockels in Germany, Devaux and Bancelin in France, and owes most to the subsequent work of Langmuir and Harkins in the United States, N. K. Adam and Rideal in England, and Dervichian in France. However, the first record is that of Benjamin Franklin in 1765 who spread olive oil on a pond at Clapham Common, London, and found that it spread to a thickness of one ten-millionth of an inch, or 25 Å.

Miss Pockels introduced the essential technique in studying liquid surfaces, a technique of first sweeping them with a solid barrier such as one of fused silica, chromium plated brass, paper toweling, etc. If there is no leak around the barrier, this sweeping effectively removes all insoluble films, leaving a clean mobile liquid surface. Dusting the surface with talc or sulfur enables one to test its mobility by gentle blowing and also reveals any spreading of substances applied to the surface.

No spreading is observed with substances like sodium chloride and sucrose which dissolve and raise the surface tension. No spreading is observed with a pure liquid paraffin (Hardy 1912), but if a fatty acid or similar polar compound is first added to the paraffin, it spreads out immediately to form a very thin film. Insoluble fatty acids spread out at the rate of about a meter per second to form a permanent film in which, from the amount required to cover a given surface, it can be seen that the film is one molecule deep. The spreading obviously is due to the polar or soluble groups evidently in contact with the water. Hardy was the first to deduce such orientation of molecules in a surface, which deduction was emphasized by Reychler, Harkins, and Langmuir.

Hardy's statement was: "If the stray field of a molecule, that is of a complex of these atomic systems, be unsymmetrical, the surface layer of fluids and solids, which are close packed states of matter, must differ from the interior mass in the orientation of the axes of the fields with respect to the normal to the surface, and so form a skin at the surface of a pure substance having all the molecules oriented in the same way instead of purely in random ways. The result would be the polarization of the surface, and the surfaces of two different fluids would attract or repel one another according to the sign of their surfaces."[75]

Rapid spreading is observed with all substances that lower the surface tension, such as camphor, phenol, or soap, but they dissolve and any displaced talc comes together again. Minute impurities in the water can appreciably affect the properties of a surface film. Thus with stearic acid spread on water the presence of 1 part of aluminum in 2 billion parts of water, or 1 part of copper in 300 million parts of water can be detected.[76] "Even pure silver or gold brought into contact with distilled water gives a contamination which is detectable within a few minutes by the change in the viscosity of the stearic acid mono-layer."

An example of the way in which spreading in force-area curves may be utilized to determine the state and configuration of molecules in a surface film is the study of films of polyorganosiloxanes on water.[77]

If the surface tension of water is lowered by 30 dynes through application of palmitic acid, for example, the spreading is rapid and forceful, and the 30 dynes is usually referred to as the spreading pressure. Since it is exerted upon a film less than 25 Å deep, it is equivalent to a compression of 100 atmospheres in the two-dimensional film. In these films the principles of orientation and of like-to-like are directly apparent.,

Miss Pockels [78] compared the surface tensions of uncontaminated water and of the film-covered surface by separate measurements of surface tension. Langmuir measured instead the pressure upon a floating barrier between them. Adam refined the film balance. Wilson and McBain introduced a flexible diaphragm below the floating barrier so that any solution could be directly compared with any other or with pure water. This is called the PLAWM trough [79] from the initials of those just mentioned. In the film balance, as in the PLAWM trough, the surface area is easily varied by moving the barrier away from or toward the floating surface upon which the pressure is exerted. The Puddington balance is the simplest and the most versatile and is placed on the surface only after it has been swept clean.[80]

The two-dimensional films on water, on solutions, and even on solids have been variously classified by the different workers in this subject. They may be rigid, and therefore described as a two-dimensional solid, or simple organic substances corresponding to an area per molecule of 20.5 Å, which is taken as the cross section of the paraffin chain. However, with some derivatives such condensed or compressed films may be liquid, and some films that under high compression occupy 20.5 square Ångström units may require about 25 or 26 under somewhat less compression. The explanation first put forward by Adam [81]

is of close-packed paraffin chains and close-packed polar heads, respectively.

On less compression or upon warming, the two-dimensional liquid may approximately double in area, producing the "expanded liquid" film first observed by Labrouste.[82] Other still more extended films are recognized, until finally at very low pressures the film approaches the behavior of a two-dimensional gas in which ideally

$$FA = kT$$

where F is the spreading pressure, A the area per molecule, k the gas constant 1.372 per molecule, and T the absolute temperature, in analogy with the well known three-dimensional gas formula

$$PV = RT$$

In practice it is necessary to modify this equation to allow for departures from ideal gas behavior. For details as to the different varieties of two-dimensional films now postulated by various investigators, reference may be made to publications of Langmuir,[83] of Harkins and Boyd,[84] of Harkins and Jura,[85] of Dervichian,[86] and most recently of Ross and Secoy.[87]

For surface films on liquid and solids Ross and Secoy propose the equation

$$\left(\pi + \frac{\alpha}{\sigma - \beta}\right)(\sigma - \beta) = kT$$

where π is the spreading pressure, β is the incompressible area of each molecule; σ is the total area of the surface per sorbed molecule; α is a constant independent of temperature obtainable by a graphical method; k is the usual gas constant per molecule (R/N), and T is the absolute temperature.

Reaction with and Penetration of Surface Films

The oxidation of a film of oleic acid spread on a dilute solution of potassium permanganate is stopped by reducing the surface area and crowding the oleic acid molecules closer together so that their double bonds are inaccessible to the permanganate.[88] Similarly, if a film of lecithin is undergoing lysis by snake venom, the rate is cut 450-fold if the area allowed for each molecule of lecithin is halved. It is clear that phenomena especially interesting to biologists occur in such systems which behave so differently from matter in bulk or in non-structural phases. There is a close relation between this field [89] and the

subject of lysis, sensitization, agglutination, immunity and lethal actions, carcinogenic activity,[90] penetration of drugs, bactericidal effects of detergents, effects of oestrogenic compounds, and the nature of living membranes.

There are two kinds of interaction between films and underlying solutes. First, there is the formation of a "complex film" with two superimposed monolayers, as in the case of cholesterol with saponin or with hexadecylamine or with sodium hexadecyl sulfate. Formation of a complex film does not greatly alter the surface tension but is revealed by its large effect on surface potential.

Second is "inter penetration," which may follow formation of a complex film, or such a mixed film may be formed directly. This greatly alters the surface tension, surface potential being less affected. A monolayer can reversibly take up an organic vapor to form a mixed film.[91] Mixed films are often very stable. The formation of strongly stable mixed films can greatly reduce the physiological activity of the components.

Compression of a mixed film can squeeze out one of the constituents, rejecting it in the form of lenses, or forming a complex film, or dispersing the whole film. Compression may likewise prevent the formation of a mixed film, as in the case of lauryl alcohol, which does not enter a film of oleic acid at pH 2, although the tendency is much greater to enter a film of sodium oleate at pH 7.2.

A mixed film greatly affects transit of substances related to or soluble in either constituent. By the use of specific penetrants, the nature of a membrane or of a film can be determined. Thus the cilia of *Mytilus* are lipoidal; those of *paramecium* are proteinaceous, whereas mixed films could be both.

Spreading of One Liquid upon Another

Clerk Maxwell in the article "Capillarity" in the 1888 edition of the *Encyclopaedia Britannica* pointed out the conditions for spreading of one liquid upon another: that the surface tension of the second liquid must be greater than the sum of the surface tension of the spreading liquid and the interfacial tension between the spreading liquid and the first liquid

$$T_a > T_b + T_{ab}$$

Harkins and Feldman in 1921 transformed this into the essentially identical expression

$$S = W_a - W_c$$

where W_c is the work of surface cohesion, W_a the work of adhesion, and S the spreading coefficient, which must be positive to produce spreading. It is evident that because liquid b spreads upon a, it is not at all a necessity that a spreads upon b. Almost all organic liquids spread upon water, although water spreads upon very few organic liquids.

Multilayers Built up from Monolayers

It is easy to transfer surface films from water to a solid surface.[92] Lord Rayleigh was the first to show that a single monolayer of oily substance on glass or porcelain is enough to decrease static friction to low values. If clean glass is thrust into water covered by a monolayer of oleic acid, the glass remains clean, but when it is gradually raised out of the water it withdraws an equal area of monolayer, as can be seen from the movement of dust on the surface toward the glass. Such a film is so firmly held that it is not readily removed by washing the glass with pure benzene. Miss Blodgett uses a two-dimensional as a "piston oil" separated from the insoluble film to be transferred by a movable barrier such as a waxed silk thread. Oleic acid, as piston oil, provides a spreading force of about 29 dynes/cm; castor oil, 16 dynes/cm. The subtraction of the monolayer from the solid surface is rendered visible by the movement of the thread over an equal area. Under such conditions, and especially if there is a small amount of divalent cation available, a second monolayer is deposited when the movement of the solid surface through the film is reversed. The first film has its polar groups attached to the solid. The second film is inside out with its polar or water-wettable or hydrophilic groups exposed, the hydrocarbon portions resting upon the hydrocarbon portions of the surface layer. Thus, on each excursion in and out alternate layers are deposited upon the solid and a built-up film of a known number of layers can be placed upon the solid. Films containing over 3000 layers have been built of barium-copper stearate or chromium plated brass. By optical examination the thickness per layer and the refractive index can be ascertained with great accuracy. Under certain conditions it is possible with the naked eye to see the presence or absence of a single monomolecular layer upon such a built-up film.

Mixed films of barium stearate and of stearic acid may be made and deposited as multilayers and then the stearic acid dissolved out by a mixture of benzene or alcohol. The barium stearate remains as a skeleton containing voids in place of the stearic acid molecules which

have been removed. Hydrocarbons can fill up these voids, but the excess of hydrocarbon draws away with a high contact angle. The hydrocarbon may again be extracted, leaving a skeleton. Wetting agents can render the pores of the skeleton hydrophilic and permeable to water and salt solutions. Barium stearate multilayers are both hydrophobic and oleophobic, that is, nonwettable by either water or oil. Zisman and coworkers [93] have shown how to produce by adsorption monolayers that are both hydrophobic and oleophobic. In explaining properties of built-up multilayers it is often necessary to assume that the outside layer can be turned inside out by contact with specific reagents. The thickness of a monomicellar film of colloidal ferric hydroxide has been shown by Langmuir-Blodgett technique to be 30–40 Å, and it could be shown that the particles were platelike, as has been shown by so many other methods.[94]

There are many industrial applications of adsorption and wetting of solids by liquids. These include film lubrication, use of horticultural sprays and insecticides, bleaching, deodorizing, decolorizing, purification, detergency, wetting of textiles, dyeing, tanning, use of adhesives, and flotation.

The subject matter of this chapter has been most extensively reviewed in *The Sorption of Gases and Vapours by Solids*, by J. W. McBain (Routledge, London, 1931, 6000 references); the encyclopaedic *Kapillarchemie* by H. Freundlich (Band I, 1930, Band II, 1932, Akademische Verlagsgesellschaft, Leipzig); and especially for films by N. K. Adam in *The Physics and Chemistry of Surfaces* (Oxford University Press, London, 1941).

REFERENCES

Text References

1. It was for this reason that the author introduced the term "sorption" replacing the narrower specialized terms absorption, adsorption, etc.; justification for this may be found in the following pages.
2. FROSCH, C. J. and HAUSER, E. A., *Ind. Eng. Chem.*, (Anal. Ed.), **8**, 423 (1936)
3. DEVAUX, H., *J. Physique*, 3(4), 450 (1904); 2(5), 699 (1912)
4. DEVAUX, H., *Actualités Scientifiques et Industrielles*, **932**, 39 (1942)
5. AFANASIEV, P. V., TALMUD, B. A. and TALMUD, D. L., *Acta Physicochim.* (USSR), **5**, 853 (1936)
6. SHEPPARD, S. E., *Science*, **100**, 545 (1944)
7. PARKHURST, K. G. A., *Chem. and Ind.*, (Jan. 15, 1949) p. 44
8. MCBAIN, J. W., *The Sorption of Gases and Vapours by Solids*, George Routledge and Sons, Ltd., London (1931)
   SMITHELLS, C. J., *Gases and Metals*, John Wiley & Sons, New York (1937)
   BRUNAUER, S., *The Adsorption of Gases and Vapors*, Princeton University Press (1943)

9. MCBAIN, J. W., JACKMAN, D. N., BAKR, H. M. and SMITH, H. G., *J. Phys. Chem.*, **34**, 1439 (1930)
   MCBAIN, J. W., LUCAS, H. P. and CHAPMAN, P. F., *J. Am. Chem. Soc.*, **52**, 2668 (1930)
10. EMMETT, P. H. and BRUNAUER, S., *J. Am. Chem. Soc.*, **59**, 1553, 2682 (1937);
   BRUNAUER, S., EMMETT, P. H. and TELLER, E., *ibid.*, **60**, 309 (1938)
   HARVEY, E. N., *Symposium on Paint and Paint Materials*, A.S.T.M., Philadelphia (1947) pp. 60–84
11. HARKINS, W. D. and JURA, G., *J. Am. Chem. Soc.*, **66**, 1362, 1366 (1944)
12. WHITE, L., *J. Phys. Coll. Chem.*, **51**, 644 (1947)
13. ROSS, S., *J. Phys. Coll. Chem.* **53**, 583 (1949)
14. DOLE, M., *J. Chem. Phys.*, **16**, 25 (1948)
15. DOLE, *loc. cit.*
16. MCBAIN, J. W. and DUNN, R. C., *J. Coll. Sci.*, **3**, 303 (1948)
17. SCHOFIELD, R. K., *Nature*, **160**, 408 (1947)
18. REHBINDER, P., *Z. physik.*, **72**, 91 (1931); *Reduction of Hardness in Drilling* (Russian), Acad. Sci. USSR, Moscow (1944), pp. 1–200
19. ENGELHARDT, W. V., *Naturwissen.*, **33**, 195 (1946)
20. TANNER, H. G., *J. Am. Chem. Soc.*, **45**, 437 (1923)
21. DEUTSCH, D., *Ber.*, **60**, 1036 (1927); *Z. physikal. Chem.*, **136**, 353 (1928); THIEL, Z., *Elektrochem.*, **35**, 266 (1929); see Chap. 2.
22. HITCHCOCK, D. I., *J. Gen. Physiol.*, **8**, 61 (1925)
23. MILLER, E. J., *Colloid Symp. Monog.*, V, 70 (1928)
24. BIGELOW, W. C., PICKETT, D. L. and ZISMAN, W. A., *J. Coll. Sci.*, **1**, 513 (1946)
25. MAIR, B. J., GABORIAULT, A. L., and ROSSINI, F. D., *Ind. Eng. Chem.*, **39**, 1081 (1947)
26. HIRSCHLER, A. E. and AMON, S., *Ind. Eng. Chem.*, **39**, 1585 (1947)
27. HOLMES, H. N. and MCKELVEY, J. B., *J. Phys. Chem.*, **32**, 1522 (1928)
28. CLAESSON, I. and CLAESSON, S., *Arkiv Kemi, Mineral. Geol.*, 19A(5) (1944)
29. KISELEV, A. V., VORMS, I. A., KISELEVA, V. V., and SHTOKVISH, N. A., *J. Phys. Chem.* (USSR), **19**, 83 (1945)
30. HUTCHINSON, E., *J. Coll. Sci.*, **3**, 219 (1948)
31. BAKR, A. M. and KING, J. E., *J. Chem. Soc.*, **119**, 454 (1921)
32. MICHAELIS, L. and RONA, P., *Biochem. Z.*, **103**, 19 (1920)
33. BARTELL, F. E. and MILLER, E. J., *J. Phys. Chem.*, **28**, 992 (1924)
34. MCBAIN, J. W., *The Sorption of Gases and Vapours by Solids*, G. Routledge & Sons, Ltd., London (1931) pp. 122–123, many references
35. COHAN, L. H., *Chem. & Eng. News*, **23**, 2082 (1945)
36. BACH, N., *Acta Physicochim.*, **14**, 463 (1941)
37. NACHOD, F. C., *Ion Exchange, Theory and Application*, Academic Press, New York (1949)
   DUNCAN, J. J. and LISTER, B. A. J., *Quarterly Reviews*, 2(4), 307–348 (1948)
   For lecture demonstration of ion exchange see WALTON, H. F., *J. Chem. Ed.*, **23**, 454 (1946)
38. WAY, THOMAS, *J. Royal Agr. Society*, **11**, 313 (1850)
39. GEDROIZ, K. K., *Kolloidchem. Beihefte*, **29**, 149 (1929)
40. CHURCH, H. F., *J. Soc. Chem. Ind.*, **66**, 221 (1947)
41. MCBAIN, J. W., LEE, W. W., MERRILL, R. C., and O'CONNOR, J. J., *Chem. Prods.* (London), **4**, 19 (1941)
42. HOROWITZ, K. and PANETH, F., *Z. physikal. Chem.*, **89**, 513 (1915)
43. ADAMS, B. A. and HOLMES, E. L., *J. Soc. Chem. Ind.*, **54**, 1 (1935)
44. A brief review of the status of industrial application of sorption is given by the Editorial Staff in *Chemical Industries*, **61**, 625 (1947), and another which also summarized the state of the theory is given by R. KUNIN, *Ind. Eng. Chem.*, **40**, 41 (1948)

45. PANETH, F., *Physik. Z.*, 15, 924 (1914)
46. STRAIN, H. H., *Chromatographic Adsorption Analysis*, Interscience Publishers, New York (1942)
    ZECHMEISTER, L. and CHOLNOKY, L., *Principles and Practice of Chromatography*, Chapman and Hall, London (1941)
    WILLIAMS, TREVOR I., *An Introduction to Chromatography*, Chemical Publishing Co., Brooklyn (1947)
    CASSIDY, HAROLD G., *et al.*, "Chromatography," Vol. 49, Article 2, *Annals of The New York Academy of Sciences* (1948)
    Introductory Notes on Chromatography, The British Drug Houses, Ltd., Poole, England (1948)
47. WILSON, J. N., *J. Am. Chem. Soc.*, 62, 1583 (1940)
48. GLUECKAUF, E., *Proc. Roy. Soc.*, 186A, 35 (1946); *J. Chem. Soc.*, 1302 (1947); COATES, J. I. and GLUECKAUF, E., *ibid.*, 1308 (1947); GLUECKAUF, E. and COATES, J. I., *ibid.*, 1315 (1947)
49. CONSDEN, R., GORDON, A. H. and MARTIN, A. J. P., *Biochem. J.*, 38, 224 (1944); CONSDEN, R., *Nature*, 162, 359 (1948) (with 51 references)
50. WILLIAMS, R. J. and KIRBY, H., *Science*, 107, 481 (1948); PARSONS, D. S., MORRISON, G. A., and FISHER, R. B., *Nature*, 161, 764 (1948)
51. BROWN, F., *et al.*, *Nature*, 161, 720 (1948)
52. LEDERER, M., *Nature*, 162, 777 (1948)
53. GOPPELSROEDER, F., *Kolloid-Z.*, 4, 23 (references to previous very extensive publications), 94, 107, 191, 236, 312 (1909); 5, 52, 109, 159, 200, 250, 305 (1910); 6, 42, 111, 174, 213, 268 (1910)
54. JONES, GRINNELL and RAY, W. A., *J. Am. Chem. Soc.*, 57, 957 (1935)
55. HARKINS, W. D. and JORDAN, H. F., *J. Am. Chem. Soc.*, 52, 1751 (1930); CUPPLES, H. L., *J. Phys. Coll. Chem.*, 51, 1341 (1947)
56. WILHELMY, L., *Pogg. Ann.*, 119, 177 (1863); 122, 1 (1864)
57. QUINCKE, G., *Pogg. Ann.*, 139, 1 (1870)
58. GOUY, G., *Ann. Physik*, 6, 5 (1916)
59. MACK, G. L. and BARTELL, F. E., *J. Am. Chem. Soc.*, 54, 936 (1932)
60. HAUSER, E., *J. Phys. Chem.*, 40, 973 (1936); FORDHAM, S., *Proc. Roy. Soc.*, London, 194A, 1 (1948)
61. FERGUSON, A. and DAWSON, P. E., *Trans. Farad. Soc.*, 17, 384 (1922)
62. JAEGER, F. M., *Z. anorg. Chem.*, 101, 1 (1917); SUGDEN, S., *J. Chem. Soc.*, 121, 858 (1922)
63. HARKINS, W. D. and BROWN, F. E., *J. Am. Chem. Soc.*, 41, 499 (1919)
64. LORD RAYLEIGH, *Proc. Roy. Soc.*, 34A, 130 (1882); LENARD, *Wied. Annalen*, N. F. 30, 209 (1887); SITZBER, *Heidelberg Akad.*, 5A, 28 (1914); ADDISON, *J. Chem. Soc.*, 98 (1945)
65. BUCHWALD, E. and KÖNIG, H., *Ann. Phys.*, 26, 659 (1936)
66. LORD RAYLEIGH, *Proc. Roy. Soc.*, 47A, 281 (1890)
67. GARNER, F. H. and NISSAN, A. H., *Nature*, 158, 634 (1946)
68. MCBAIN, J. W. and HUMPHREYS, C., *J. Phys. Chem.*, 36, 300 (1932); MCBAIN, J. W. and WOOD, L. A., *Proc. Roy. Soc.*, 174A, 286 (1940)
69. MCBAIN, J. W., FORD, T. F. and WILSON, D. A., *Kolloid-Z.*, 78, 1 (1937)
70. MILES, G. D. and SHEDLOVSKY, L., *J. Phys. Chem.*, 48, 57 (1944)
71. BRADY, A. P., *J. Phys. Colloid Chem.*, 53, 56 (1949)
72. MCBAIN, J. W. and PERRY, L. H., *Ind. Eng. Chem.*, 31, 35 (1939)
73. For references see DOSS, K. S. G., *Kolloid-Z.*, 86, 205 (1939)
74. SOBOTKA, H., "Monomolecular Layers. Their Application in Physiology and Medicine," in *Medical Physics*, Otto Glaser, ed., Year Book Publishers, Inc., Chicago (1944)

TRURNIT, H. J., "Monomolecular Films," *Fortschritte der Chemie Organischer Naturstoffe*, Vol. IV, Springer Verlag, Vienna (1945)

75. HARDY, W. B., *Proc. Roy. Soc.*, **86A**, 634 (1912)
76. LANGMUIR, I. and SCHAEFER, V. J., *J. Am. Chem. Soc.*, **59**, 2400 (1937)
77. FOX, H. W., TAYLOR, P. W., and ZISMAN, W. A., *Ind. Eng. Chem.*, **39**, 1401 (1947)
78. POCKELS, A. A., *Naturwissenschaften*, **9**, 137, 149 (1917)
79. MCBAIN and WILSON, *loc. cit.*
80. PUDDINGTON, J. E., *J. Coll. Sci.*, **7**, 505 (1946)
81. ADAM, N. K., *J. Phys. Chem.*, **29**, 87 (1925)
82. LABROUSTE, H., *Ann. Physik*, **14**, 164 (1920)
83. LANGMUIR, I., *J. Chem. Phys.*, **1**, 756 (1933)
84. HARKINS, W. D. and BOYD, E., *J. Phys. Chem.*, **45**, 20 (1941)
85. JURA, G. and HARKINS, W. D., *J. Am. Chem. Soc.*, **68**, 1941 (1946)
86. DERVICHIAN, D. G., *J. phys. radium*, **10**, 333 (1939); *J. Chem. Phys.*, **7**, 931 (1939)
87. ROSS, S. and SECOY, C. H., *J. Phys. Colloid Chem.*, **53**, 306 (1949)
88. HUGHES, A. H. and RIDEAL, E. K., *Proc. Roy. Soc.*, **140A**, 253 (1933); SCHULMAN, J. H. and RIDEAL, E. K., *ibid.*, **122B**, 29, 46 (1937)
89. DAVSON, H. and DANIELLI, J. F., *The Permeability of Natural Membranes*, Cambridge Univ. Press (1943); RIDEAL, E. K. and SCHULMAN, J. H., *Nature*, **144**, 100, 693 (1939)
90. CLOWES, G. H. A., *Surface Chemistry*, Publication #21 of the A.A.A.S. (1943), pp. 1–16
91. DEAN, R. B. and MCBAIN, J. W., *J. Colloid Sci.*, **2**, 383 (1947)
92. BLODGETT, K. B., *J. Am. Chem. Soc.*, **57**, 1007 (1935); BLODGETT, K. B. and LANGMUIR, I., *Phys. Rev.*, **51**, 946 (1937); LANGMUIR, I., *Proc. Roy. Soc.*, **170A**, 1–39 (1939)
93. BIGELOW, W. C., PICKETT, D. L. and ZISMAN, W. A., *J. Coll. Sci.*, **1**, 513 (1946); also *J. Chem. Phys.*, **9**, 534, 729, 789 (1941)
94. MOKRUSHIN, S. G., *Trans. Farad. Soc.*, **43**, 1 (1947)

CHAPTER *4*

THE EFFECTIVE DEPTH OF SURFACES

One of the fundamental problems of science is to establish the effective ranges of molecular attractions. The classical physicists of the 19th century considered that there were direct forces of attraction over distances which amounted to thousands or perhaps hundreds of thousands of Ångström units, that is, up to many microns.

During the first decades of the 20th century, however, the recognition of the electrical structure of matter influenced scientific opinion to assume, without new evidence, that the direct range of molecular attraction amounted to only a few Ångström units; and an extreme view, under the influence of Langmuir, held that not adjacent molecules but only their adjacent atoms had any important influence upon each other. A third point of view recognizes the short range of direct attraction but considers that it must be relayed from molecule to neighboring molecule through impressive distances. The conflicting consequences of these beliefs are of the greatest importance.

Is it true that the properties of a surface are conditioned solely by the exposed atoms or molecules and that underlying layers are unaffected? Or, on the contrary, is it true that a liquid in the neighborhood of its surface and to a great depth differs markedly from bulk liquid? Is the surface region of a liquid to a depth of thousands of Ångström units often semirigid? Is it due to relayed action that the nucleus of a living cell conditions the contents of a cell for a definite but limited range? In a living organism is most of the liquid under the influence of neighboring surfaces and colloidal particles? Are there significant effects of deep surfaces in adhesive joints, in lubrication, in heat transfer, in all thin films, in the x-ray structure of soaps, clay, and proteins, in all surface effects, and in the neighborhood of all

colloidal particles? Is the reinforcement of a rubber tire by carbon black due in part to effects relayed through the rubber by the solid particles? What laws must geologists and petroleum engineers assume in calculating the flow of oil and brine through finely porous structures? In all these cases we find that the actual behavior is that which would exist if the surface zone were deep. Is this always due to dust, rough surfaces, contamination, and experimental error?

Direct and Relayed Forces of Attraction

Before considering the evidence it should be stated that no one doubts the existence of long range forces where free charges or charged surfaces are concerned. In addition, the Dutch school [1] favors the idea that the residual effects of van der Waals-London attraction might be felt at a distance of some thousands of Ångström units where large or extended surfaces are concerned, although the attraction between individual pairs of molecules across that distance is completely negligible. However the potential energy between one sq cm of a surface and another sq cm of a surface opposite falls off only with the inverse square of the distance between them (Debye 1948). Finally, it should be made quite clear that, apart from such views, the only mechanism that appeals at the present day is that of relayed action of intrinsically short-range forces in the same way that a magnet can pick up a long chain of iron filings to a distance far beyond that at which the direct attraction has become negligible. This is the definite concept introduced by Hardy in 1912.[2] Hardy's statement is as follows:

> The radius of the sphere of influence of a single molecule . . . has an extremely small value, not much greater than the diameter of the molecule itself. Quite another kind of range of molecular action is possible in close packed structures such as fluids and solids, namely a strain transmitted from molecule to molecule. . . . It is this irradiation of strain in a close packed structure which may extend so deeply as to modify the molecular state of a skin some hundreds of microns in depth; (that is, some millions of Ångströms in depth).

The liquid state is now recognized as being closely akin to the solid state; and in the liquid, neighboring molecules tend to take up positions corresponding to those of the space lattice of the solid, although this order does not in general extend over many molecular diameters in ordinary bulk liquid. The x-ray examination of liquids by Sogani [3] and Stewart [4] has shown the tendency of neighboring molecules in a liquid to take up parallel or mutually oriented positions, particularly if the molecules are elongated or otherwise anisotropic; this phenomenon was given the name "cybotaxis."

Bulk liquids are then close-packed with some tendency toward orderly arrangement. This tendency might be strongly reinforced and developed by proximity to a polarized surface or layer of oriented molecules. An example of such a surface might be an ionic crystal or a film of an insoluble fatty acid whose dipoles are all parallel to each other and approximately normal to the interface. Sir William Hardy [5] and McBain [6] independently suggested a relayed action of this sort extending inwards in molecular chains of solute molecules from an oriented adsorbed film on the surface, but the view had already been clearly expressed by Hardy in his remarkable paper of 1912, which still repays careful study. In 1942 [7] McBain from consideration of the stability of oil foam postulated that solvent as well as solute molecules would likewise be so oriented.

Calculations such as those of Lennard-Jones and Miss Dent [8] show that the direct range of action of a crystal surface without special orientation is not insensible at a distance of 10 Å, where a potassium chloride crystal can increase the density of argon at room temperature two- or threefold. How important this can be when already close-packed molecules merely have to be oriented, and thus relay the influence, is clearly evident. Likewise, the effect upon a vapor that is already nearly saturated would be to produce thick multilayer adsorption. Indeed, the so-called B.E.T. method of determining surface areas rests upon the now commonly accepted assumption of such multilayer adsorption merging into capillary condensation. If then, nearly saturated liquid produces multilayer adsorption, completely saturated vapor or the liquid itself must produce a deep surface zone.

Lennard-Jones and Dent [9] add the statement:

> It may be of interest to observe that if a number of long dipoles were arranged on the crystal surface opposite each lattice point and oriented alternately in one direction and then the other, the electrostatic field outside the projecting ends of the dipoles would be approximately the same as that of the crystal itself, however long the dipole. (This is because the first layer of crystal has by far the most important effect, cf. equations (2.06 and 2.07)). It would then be easy for another such layer of dipoles to form on top of the first and so on indefinitely. This may be relevant with regard to some phenomena of adhesion recently observed by Hardy and Nottage [10] and McBain and Lee.[11]

It is interesting to mention two fields in which almost all the workers take for granted deep, immobilized surfaces without perhaps realizing that if true for those cases it must in some degree be of general occurrence.

First, in liquid crystalline or anisotropic liquid phases the molecules are now definitely known to be strung together and ordered over

large visible tracts, as is readily seen from their double refraction of light and their parallel extinction throughout any one of these tracts. This has had but little impact on general science, because a generation ago the very existence of liquid crystals was suspect and as late as 1930 the leading investigators represented in the 347-page symposium on the subject [12] could not even agree upon a name for it. Liquid crystals commonly orient themselves upon any adjacent solid surface.

The other subject in which authorities agree in assuming that surfaces of water and of aqueous solutions are deep is the field of electrokinetics. There is no correspondence between electrokinetic potential and the known Nernst or thermodynamic potential differences between the bulk phases, and the suggestion of Freundlich is everywhere adopted that this must be because a portion of the charges of the diffuse double layer lie in a layer of water near the surface that is so solid that the charges therein do not move even under the great stresses of the electric field. The thickness of this solid or rigidified water layer was estimated by Cotton and Mouton [13] as 25,000 Ångström units. Eversole and Lahr [14] suggest that it is between 8 and 63 Å.

To explain the viscosity caused by spherical colloidal particles Hatschek [15] had to assume an envelope of rigidified liquid 9 Å thick to meet the Einstein equation, although this had already taken fully into account the adherent liquid water. Hence this would have to be completely solidified, or a correspondingly thicker layer of less completely rigidified liquid would have to be postulated.

The Importance of the Outermost Layer of Molecules

Naturally, for many purposes the outermost layer of molecules of liquid is by far the most important. However, it will first be shown that this is not the whole story. Then more definite evidence that surfaces are deep will be adduced.

The pure paraffins exhibit a wide range of surface tensions, although they all can expose only paraffin chains and their exteriors are assumed to consist entirely of methyl groups. Mrs. McBain and Perry [16] showed that the surface tension of such hydrocarbons can be appreciably lowered by dissolving in them polar compounds such as lauryl sulfonic acid. Thus Nujol is lowered from 33.5 to 24.8 dynes, hydrogenated tetraisobutylene from 25.0 to 19.4 dynes. There are three systems whose externally exposed groups are identical: pure oleic acid, a close-packed film of oleic acid on water, and a solution of sodium oleate. Nevertheless, their surface tensions are very different: 32.5, 40, and 25 dynes, respectively.

Aqueous solutions of the lower fatty acids exhibit a lowering of surface tension which, according to the adsorption theorem of Gibbs, shows that over a wide range the composition of the surface is constant, being somewhat less than a monomolecular layer of the fatty acid. Nevertheless, the surface tension of this constant layer changes with the concentration of the underlying solution.

Liquid Layers May Have an Ordered Structure

X rays provide definitive proof that layers of water and of other liquids many molecules deep can be rigorously ordered. This is true, for example, of hydrocarbons solubilized in lamellar micelles in soap solutions. (See Chap. 17.) Crystalline clays consist of alternate layers of silicate and of water. (See Chap. 26.) A thorough x-ray study of methemoglobin by Boyes-Watson, Davidson, and Perutz [17] has established that layers of hemoglobin within the crystal are separated by definite rigidified uniform layers of water, 15–25 Å thick, through which ions readily diffuse. (See Chap. 24.)

A Crystal Surface May Orient and Deeply Affect a Liquid

A liquid adjacent to a crystal surface can exhibit orientation to a visible depth. Taylor and King [18] in 1933 found that fatty acids melted on glass remain anisotropic for several degrees above their melting point and show interference colors. From these colors they estimated the thickness of oriented molecules to be five or six thousand Ångström units. Brummage [19] found by electron diffraction that orientation persists for a number of liquids spread on metals at temperatures well above their melting points. Bangham [20] produced visibly thick films of benzene on mica which were not wetted by ordinary benzene in bulk, the latter drawing away as lenses with a definite contact angle. Langmuir also found that built-up skeleton multilayers of barium stearate, whose pores had been filled with a hydrocarbon, were not wetted by excess of the same hydrocarbon. Similarly Terzaghi [21] found that water does not freely wet moist clay.

Hardy in 1925 referred to x-ray examinations by Bragg showing that films of grease on solid surfaces were similarly oriented, but his reference does not make it clear that the observations were made above the melting point of the grease. The same comment applies to the thick, visibly oriented films of grease used for sealing desiccators to which Kumler [22] has drawn attention.

Clark, Lincoln, and Sterrett [23] x-rayed films of lubricating oil containing polar compounds spread on metals, and from the sharpness of the x-ray lines they estimated that the polar additives are oriented to a depth of 6000 Å, but their work may refer to frozen samples. Trillat [24] apparently obtained similar results with liquid, but his results may be an artifact because of his having employed a slit collimating system (S. S. Marsden, Jr.).

Derjaguin's Proofs of Deep Surfaces of Liquids

Derjaguin, more than anyone except Hardy, has given long and careful attention to the determination of the effect of liquid surfaces. He and his collaborators have presented proof that a liquid film in contact with a surface of glass or mica is rigidified to a depth of the order of 1000 or 1500 Å.[25] They found extreme rigidity in films 350 Å deep. The most recent ingenious device used in the Moscow laboratory for surface forces is the "blow-off" method [26] in which a shearing force on a liquid in a narrow wedge-shaped or U-shaped slit is produced by blowing a stream of air along its surface. The motion of the liquid is followed by photographing interference bands produced in it and examining their contours. This method demonstrated in certain cases the existence of a yield value in the surface layer as well as deviations from Newtonian viscosity.

Difficulty of Interpretation of Other Results Obtained with Liquids between Two Solid Surfaces Simulating Behavior of Deep Surfaces

Liquids between two solid surfaces or in capillaries and crevices behave as if the surface zone was rigidified in depth. However, the technique of experiments with solid surfaces is difficult because of the small dimensions involved and because all solids are elastic and, hence, do not preserve definite dimensions under varying stress. For example, a steel cylinder can readily be inserted into a hole in a steel block at least 0.0001 inch smaller than the cylinder; this latter must be lubricated and kept in continuous smooth motion if violent seizure is to be avoided.

It is often possible to assume that ultramicroscopic asperities exist upon the solid surface and that these projections on the solid surface touch or interlock and prevent enclosed liquid films from thinning or from showing the usual bulk fluidity. Alternatively, it is often suggested that invisible dust particles have settled upon the solid surface

and have such properties as to explain the phenomena commonly observed. It is sometimes difficult to apply these assumptions consistently, but while recognizing the behavior [27] referred to, simulating or perhaps sometimes actually due to thick films, it is possible to ignore all such experiments in looking for theoretical significance. Such surfaces need examination by the electron microscope to prove or disprove the presence of hypothetical asperities.

An exceptionally careful study is that of Bulkley [28] in 1931 which he summarizes in the following words:

> Measurements of the flow of various fatty and mineral oils through steel and nickeline capillary tubes as small as 0.18 and 0.13 mm inside radius, respectively, have shown that there is no clogging of the tubes over long periods of time when the liquids are adequately filtered. Similar measurements have been made on a much greater variety of liquids using platinum and glass capillaries as small as 9.35 μ and 5.59 μ inside radius, respectively. By applying to these measurements certain criteria based on the laws of simple viscous flow it has been shown that these liquids retain their ordinary bulk viscosity at least as close to the solid surfaces as 200 or 300 Å.

However, reference to the actual records of his experimental work shows the almost insuperable difficulty of carrying out measurements significant or accurate to such dimensions.

Bowden's experiments [29] lead him to believe that any completely rigidified liquid must be less than 1000 Å thick. Derjaguin [30] pointed out that this did not conflict with his finding that the partially rigidified liquid extended out about 2000 Å, and, indeed, that Bowden's experiments could lead to the same numerical result.

Fragility of Deep Surfaces

Especially should it be remembered that there is little or no evidence for the existence of very thick, oriented layers in the neighborhood of surfaces which are perfectly fresh or are in rapid motion. Thick films are delicate and are readily thinned and progressively swept away, down to a matter of a few hundred Ångström units.

Evidence for Depth Where Solids Are Not Involved

The preceding argument against experiments employing solid surfaces cannot be adduced when only one is involved, as where a bubble is pressed against a solid and the intervening film of liquid is found to remain thick and rigid.[31] Still more is this the case when no solids are involved, as in Hardy's film of water between air and a layer

of castor oil only 180 Å thick, where the range of action in the castor oil had to be at least half this value; or again in Hardy's stable film of water between two monomolecular layers of oleic acid,[32] or in the similar one in soap bubbles.[33] Similarly, this holds true in the experiments of Hazlehurst and Neville,[34] entitled "Cybotaxis in Surfaces."

Bubbles of air rising in doped oil [35] behave as if surrounded by a thick film. The observed, although limited, stability of various bubbles cannot be explained if the liquid film has merely the properties of the liquid in bulk. Neither can the stability of films of oils 300 to 5000 Å thick that support a steel ball and separate it from contact with mercury.[36]

Inconclusiveness of Evidence from Adsorption

A further very large body of evidence is to the effect that sorption on solid surfaces is often great enough to cover the visible surface with a film very many molecules thick.[37] The evidence can be provisionally held in abeyance with the verdict "Not Proved," because it can almost always be conceived that the surface possessed minute cracks or pores, that it was covered by porous products of corrosion, or that in some way the actual surface was far greater than the visible surface and therefore the layer might be correspondingly thinner. Even so, Joris and Taylor [38] estimate that water is sorbed on glass 16–62 molecules deep. Nevertheless, the now commonly accepted B.E.T. method presupposes deep layers — the subject of this chapter.

Composite Adhesive Joints

It is a remarkable fact that in many cases the strength of a cemented joint between two solids depends upon the nature of both solid surfaces as well as upon the cement. Highly polished as well as original crystal surfaces can be strongly glued together by pure chemicals. The thinner the film, the stronger the joint, up to a value many times the strength of the film itself as ordinarily measured. This is true even of soldered joints. The behavior is as if the influence of the solid extends through even a microscopically visible depth of adhesive.

Hardy and Nottage,[39] as well as McBain and his collaborators,[40] have found many examples of this behavior. One such instance [41] is that a joint made with de Khotinsky cement between two polished copper surfaces showed a tensile strength of 5110 pounds per square inch, but only 300 pounds per square inch between two tungsten

surfaces. The obvious strength of a composite joint with the same adhesive between one surface of copper and the other of tungsten would have been only 300 pounds per square inch, but it gave the intermediate value of 780 pounds per square inch although in all cases the joint broke within the adhesive layer. With metals using any non-corrosive adhesive the strength of adhesive joints always increases with the tensile strength of the metal joined, with its elasticity and with its hardness, and decreases with the atomic volume and the compressibility of the metal.

Talmud [42] found that the strength of a joint is profoundly affected and increased by inserting a monomolecular layer on one or both of the solid surfaces. Similar composite effects on boundary lubrication have been found by Hardy and Doubleday.[43]

Other Properties of Thin Films of Liquid

Many properties of thin films differ from those of the same material in bulk. Some are the electrical conductance,[44] thermal conductance, dielectric constant, density, and possibly surface tension. For example, Meyer [45] in 1920, like Gratz in 1879, found that for heat transfer between water and copper the behavior is as if a microscopically visible layer was immobilized. This effect is observable, when turbulent flow is avoided, in systems such as aqueous ethylene glycol in automobile radiators.

Rate of heterogeneous reaction as studied by Nernst and Brunner [46] leads to an approximately immobilized layer of liquid in contact with a solid, of microscopic thickness, that is, tens of thousands of Ångströms thick.

Interesting Instances of Long Range Effects

Examples of the effect of long range action are given by the work of Rothen.[47] A film of antigen on a metal can sorb layers of specific antibody as much as 300 Å deep, but appropriate enzymes destroy this ability. A layer of plastic several hundred Ångströms thick can be placed over the original layer of antigen and yet the sorption takes place. The enzyme can affect the antigen through this distance without penetrating the plastic. The interaction and distance depend on both the thickness of the antigen layer and that of the screening layer.

Bradley [48] has found that mica affects the form of crystallization of ammonium iodide and that this influence can be exerted through a

film of cellulose acetate or rubber as much as 1000 Å deep, or even through liquid octane.

Finch [49] found that the recrystallization of nickel follows that of a copper substrate up to layers 30,000 Å thick.

Conclusion

Thickened liquid surfaces may arise from the effects of an electrical double layer which can hold surfaces of like charge apart; or they may be due to ramifying aggregates of colloidal particles in loose contact, as described in Chapters 10 and 11; or, with suitable choice of liquids, they may be the result of cybotaxis at surfaces, described in this chapter. One or more of these factors may be operative in the production of tactoids or Schiller layers, of the thick films of liquid observed in certain wax emulsions, of sedimentation volumes [50] of powder which depend so much upon the liquids present, and of numerous other phenomena such as the varying adhesion of wetted particles to solids or the increased surface tension which suspended particles give to a liquid.[51] Some of these factors may be distinguished by study of dielectric constants.[52]

The considerations reviewed in this chapter have innumerable practical consequences, and they pervade the whole field of stability and behavior of colloids, for example, the settling of pigments in paints, the permeability of clays in water, and the properties of coacervates.

A more extensive survey of this field has been made by Henniker.[53]

REFERENCES

Text References

1. VERWEY, E. J. W. and OVERBEEK, J. T. G., *Trans. Farad. Soc.*, **42B**, 117 (1946); VERWEY, E. J. W., *Rec. trav. chim.*, **64**, 33; (1945) see also LONDON, F., *Surface Chemistry*, Publication #21 of the A.A.A.S., Washington (1943), p. 141
2. HARDY, W. B., *Proc. Roy. Soc.*, **86A**, 631 and 632 (1912)
3. SOGANI, C. M., *Indian J. Phys.*, **2**, 97, 113 (1927)
4. STEWART, G. W., *Phys. Rev.*, **27**, 104 (1926); **30**, 232 (1927); **31**, 1 (1928)
5. HARDY, W., *J. Gen. Physiol.*, **8**, 641 (1927)
6. MCBAIN, J. W. and DAVIES, G. P., *J. Am. Chem. Soc.*, **49**, 2230 (1927)
7. MCBAIN, J. W., N.A.C.A., Washington, D.C.
8. LENNARD-JONES, J. E. and DENT, B. M., *Proc. Roy. Soc.*, **121A**, 247 (1928); *Trans. Farad. Soc.*, **24**, 92 (1928)
9. *Loc. cit.*
10. *Proc. Roy. Soc.*, **112**, 62 (1926)
11. *Proc. Roy. Soc.*, **113**, 606 (1927)
12. *Z. Krist.*, **79**, 1–347 (1931)

13. COTTON, A. and MOUTON, H., *Les Ultramicroscopes*, Mason & Cie., Paris (1906), p. 148 (see Nordenson, *Kollchem. Beihefte*, 7, 108 (1915)

14. EVERSOLE, W. G. and LAHR, P. H., *J. Chem. Phys.*, 9, 530 (1941)

15. HATSCHEK, E., *Kolloid-Z.*, 11, 280 (1912)

16. MCBAIN, M. E. L. and PERRY, L. H., *J. Am. Chem. Soc.*, 62, 989 (1940)

17. BOYES-WATSON, J., DAVIDSON, E., and PERUTZ, M. F., *Proc. Roy. Soc.*, 191A, 83 (1947)

18. TAYLOR, A. M. and KING, A., *Nature*, 132, 64 (1933)

19. BRUMMAGE, K. G., *Proc. Roy. Soc.*, 191A, 243 (1947)

20. BANGHAM, D. H., *J. Chem. Phys.*, 14, 352 (1946); BANGHAM, *et al.*, *Nature*, 140, 237 (1937)

21. TERZAGHI, *Phys. Rev.*, 16(2), 54 (1920)

22. KUMLER, W. D., *J. Phys. Chem.*, 44, 612 (1940)

23. CLARK, G. L., LINCOLN, B. H. and STERRETT, R. R., Am. Petr. Inst., Los Angeles, Calif., Nov. 13, 1935

24. TRILLAT, J. J., *Ann. phys.*, 10, 5 (1926)

25. DERJAGUIN, B., *Nature*, 138, 330 (1936)

26. DERJAGUIN, D., KUSAKOV, M., *et al.*, *Acta Physicochim.* (USSR), 19, 541 (1944); 20, 35, 47, 548 (1945); 22, 289 (1947); *J. Exptl. Theoret. Phys.* (USSR), 16, 171, 179 (1946)

27. TERZAGHI, *loc. cit.*
    BERDENNIKOV, V. P., BRESLER, S. E., ZELMANOV, I., and STRAUF, E. A., *Phys. Z. Sowjetun.*, 4, 873 (1933)
    EVERSOLE, W. G. and LAHR, P. H., *Proc. Iowa Acad. Sci.*, 48, 275 (1941); *J. Chem. Phys.*, 9, 686 (1941)
    MARCELIN, A., *Compt. rend.*, 207, 777 (1938); *Mécanique*, 26, 35 (1942)
    NORTON, F. H. and JOHNSON, A. L., *J. Am. Ceramic Soc.*, 27, 77 (1944)
    BOWDEN, F. P., *Phys. Z. Sowjetun.*, 4, 185 (1933)
    NEEDS, J. S., *Trans. Am. Soc. Mech. Engrs.*, 62, 331 (1940)
    DERJAGUIN, B. V., *Phys. Z. Sowjetun.*, 4, 431 (1933); *Z. Physik*, 84, 657 (1933)
    HARDY, W. B. and NOTTAGE, M., *Phil. Trans. Roy. Soc.*, 230A, 1 (1932)
    MACAULEY, J. M., *Nature*, 138, 587 (1936)
    DUFF, A. W. *Phil. Mag.*, 9, 685 (1905)
    WOLKOWA, Z. W., *Kolloid-Z.*, 67, 280 (1934)
    BULL, H. B. and WRONSKI, J. P., *J. Phys. Chem.*, 41, 463 (1937)
    REEKIE, J. and AIRD, J., *Nature*, 156, 367
    DERJAGUIN, B. and KRYLOV, N. A., *Akad. Nauk. USSR, Otdel. Tekh. Nauk.* (Conference on Colloids), 2, 52 (1944)
    WILSON, R. E. and BARNARD, D. P., *Ind. Eng. Chem.*, 14, 682 (1922)
    DOVER, M. V., *Ind. Eng. Chem.*, 18, 499 (1926)
    WILSDON, B. H., BONNELL, D. R. G. and NOTTAGE, M. E., *Trans. Farad. Soc.*, 31, 1304 (1935)
    NUTTING, P. G., *Science*, 97, 74 (1943); repeated and confirmed by Kaminski, A., Stanford University (1947)
    HARDY, W. B., *J. Chem. Soc.*, 127, 1217 (1925)
    MYSELS, K. J. and MCBAIN, J. W., *J. Coll. Sci.*, 3, 45 (1948)

28. BULKLEY, R., *Bur. Standards J. Research*, 6, 89 (1931)

29. BOWDEN, *loc. cit.*

30. DERJAGUIN, B., *Nature*, 138, 330 (1936)

31. DERJAGUIN, B., KUSAKOV, M. and LEBEDEVA, L., *Compt. rend. acad. sci.* (USSR), 23, 671 (1939)

32. HARDY, W. B., *J. Chem. Soc.*, 127, 1224 (1925)

33. REINOLD, A. W. and RUCKER, A. W., *Proc. Roy. Soc.*, **35A**, 149 (1883); **53A**, 394 (1893)

34. HAZLEHURST, T. H. and NEVILLE, H. A., *J. Phys. Chem.*, **44**, 592 (1940)

35. ROBINSON, J. V., *J. Phys. Coll. Chem.*, **51**, 431 (1947)

36. TAUSZ, J. and SZEKELY, P., *Erdol u. Teer.*, **9**, 331 (1933)

37. MCBAIN, J. W., *The Sorption of Gases and Vapours by Solids*, Geo. Routledge and Sons, Ltd., London (1931)

TRILLAT, J. J. and VAILLE, R., *J. chim. phys.*, **33**, 742 (1936)

38. JORIS, G. G. and TAYLOR, H. S., *J. Chem. Phys.*, **16**, 50 (1948)

39. HARDY, W. and NOTTAGE, M., *Proc. Roy. Soc.*, **112A**, 62 (1926)

40. FULLER, G. W., Ph.D. Thesis, Stanford University, 1939

41. MCBAIN, J. W. and MORRISON, J. G., Appendix to Fuller's Thesis, *loc. cit.*

42. TALMUD, D. L., *Kolloid-Z.*, **64**, 227 (1933)

43. HARDY, W. B. and DOUBLEDAY, I., *Proc. Roy. Soc.*, **100A**, 550 (1922); DOUBLEDAY, I., *J. Chem. Soc.*, **121**, 2875 (1922)

44. BRUNINGHAUS, L., *Compt. rend.*, **188**, 1386 (1929); **192**, 151 (1931); *J. phys. radium*, **1**(7), 11 (1930)

WATSON, H. E. and MENON, A. S., *Proc. Roy. Soc.*, **123A**, 185 (1929)

45. MEYER, J., *Z. physik. Chem.*, **96**, 275 (1920)

46. NERNST, W. and BRUNNER, E., *Z. physik. Chem.*, **47**, 52, 56 (1904)

47. ROTHEN, A., *J. Biol. Chem.*, **163**, 345 (1946); **167**, 299 (1947); **168**, 75 (1947); *Science*, **106**, 509 (1947)

48. BRADLEY, R. S., *Z. Krist.*, **96**, 499 (1937)

49. FINCH, G. I., *Nature*, **140**, 800 (1937)

50. VON BUZAGH, A., *Kolloid-Z.*, **47**, 370 (1929); **51**, 105, 230 (1930); **52**, 46 (1930); *Kollchem. Beihefte*, **32**, 135 (1930); FREUNDLICH, H. and JONES, A. D., *J. Phys. Chem.*, **40**, 1217 (1936)

51. ANTONOFF, G., *J. Phys. Coll. Chem.*, **52**, 969 (1948)

MCGEE, C. G., Stanford University; unpublished

52. VOET, A., *J. Phys. Coll. Chem.*, **51**, 1037 (1947)

53. HENNIKER, J. C., *Rev. Modern Phys.*, **21**, 322 (1949)

THE PREPARATION OF COLLOIDAL SOLS

Before attempting generalizations as to the properties and behavior of colloid systems, it is advisable to consider a number of typical examples of such materials and the methods by which they are produced in this condition. It is impossible here to review the innumerable known instances, and reference has to be made to the Subject Index of *Chemical Abstracts*, *Kolloid Zeitschrift*, and *Chemisches Zentralblatt*. Von Weimarn (1906–1935) prepared many hundreds of substances in the colloidal condition in a score of solvents; some of the more intriguing were colloidal ice, colloidal sugar, a jelly of sodium chloride, colloidal sulfur in all colors of the spectrum, aqueous sols of cellulose, silk, fibroin, and chitin. Some of the older examples were collected by Svedberg in 1909, and Weiser has compiled information on the elements and on the hydrous oxides and hydroxides.

We will begin instead with an instructive illustration of the nature and role of the stabilizing agent. Giles and Salmon [1] prepared sols of silver in pure stearic acid by dissolving silver stearate in molten stearic acid at 180° C and passing in a current of hydrogen. Dark red sols containing as high as 10% of free silver were obtained. If the reduction were carried to completion, no sol could be formed. A small amount of silver stearate was sufficient to act as stabilizing agent for large amounts of silver. Evidently the silver particle is linked to the stearic acid medium by molecules of oriented silver stearate.

Formation of Stable Colloids

It was emphasized in Chapter 1 that some substances are most stable in the colloidal condition and others in true solution, depending upon the solvent. Such stable systems form spontaneously by merely

bringing together the components; soaps are crystalloidal when they dissolve in alcohol, but form colloidal electrolytes in aqueous solution. Walden found for tetrapropyl- and tetraisoamyl-ammonium iodides in organic solvents that association increased with decreasing dielectric constant of the solvent, the isoamyl compound being colloidal in benzene and in carbon tetrachloride.[2] Bituminous coal dissolves in oils at 300–400° as a stable colloid, giving a sol and a jelly analogous to those of gelatin in water.[3] These "carbon" sols may be contrasted with the "carbon" sols formed by the interaction of sucrose with sulfuric acid and dilution with water.[4]

Formation of Unstable Colloids

All colloidal systems possess some measure of stability, otherwise they would not exist. How slight is the stability shows in the readiness with which many coagulate when disturbed or when exposed to slight contamination. To make them stable a sufficient quantity of stabilizing agent must be introduced or some protective colloid added.

Unstable colloids are illustrated by the pink suspension of selenium formed by throwing a true molecular solution of selenium in carbon disulfide into a large amount of ether, in which selenium is insoluble. The stabilizing agent is left to chance, but the ether may serve.

Von Weimarn assumes that usually the reaction producing the insoluble substance is instantaneous. To form a particle from such molecules it is necessary that the concentration of the reaction product exceed its true solubility. Each particle grows from an initial nucleus appearing spontaneously, or occurring, or deliberately placed in the medium. Hence, besides the stabilizing agent one must consider:

1. The rate of formation of nuclei and the number of nuclei present
2. The rate of growth of nuclei once they are present.

These two factors are largely independent, although both are influenced by the solubility and by the extent to which this is exceeded, that is, by the degree of supersaturation. Familiar examples of supersaturated liquids are honey and glycerin, in which on cool days the rate of formation of nuclei is so low that they are usually absent, whereas many other liquids can scarcely be supersaturated or undercooled without crystallization. Likewise the rate at which nuclei grow in size depends greatly upon the substance. It is readily followed by introducing a nucleus or crystal at one end of a long tube filled with the liquid and observing the linear rate of growth along the tube. This varies from 1 mm per minute for betol (β naphthyl salicylate) to 60,000 mm per minute for phosphorus.

Because of the many other influences affecting these two rates, von Weimarn's attempted numerical formulations are not quoted here. Instead, we reproduce two diagrams to illustrate his predictions as to the effect of solubility, time, and concentration of the products of the reaction upon the size of the particles produced.

From these diagrams it is readily seen that the more slowly a particle is allowed to grow, other nuclei remaining absent, the coarser it can become. This is illustrated in the method of obtaining a large crystal by allowing it to grow in a very slightly supersaturated medium. The most striking prediction from von Weimarn's

FIGURE 5·1 Variation of size of individual crystals with concentration of reacting solutions for precipitates of different solubilities. Taken from von Weimarn, P. P., *Mem. Coll. Sci. Kyoto Imp. Univ.*, **Series A, VII, 415 (1924).**

diagrams is that particles will be at their finest when formed from very dilute or from very concentrated solutions. In very dilute solutions diffusion is very slow because the concentration gradient is so small, and sufficient nuclei will usually appear to relieve the slight supersaturation locally. Hence, only a very small amount of material will be available for each particle. In extremely concentrated solutions, again, nuclei appear so quickly and in such numbers that supersaturation is immediately relieved before appreciable diffusion can occur. This behavior is well shown by a series of three experiments proceeding from saturated solutions of ferric chloride and potassium ferrocyanide. In the first, a drop of each is diluted to 100 cc and the two are mixed to give a beautiful deep blue, transparent sol of Prussian blue which passes unaltered through filter paper. In the second, 1 or 2 cc of each are diluted to 100 cc before mixing. The precipitate of Prussian blue after standing a few moments is completely held back by filter paper, the filtrate being nearly colorless or a pale green. In the third experiment, one part of saturated ferric chloride is poured into two parts (the chemically equivalent amount) of saturated potassium ferrocyanide
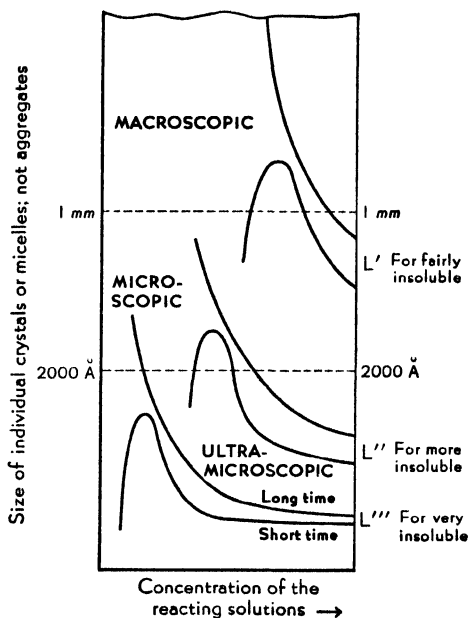
Figure on left:

Infinite time

Short time

Particle size of precipitate →

Slight   Moderate   Great supersaturation

Concentration of
reacting solutions →

FIGURE 5·2 More detailed study of a precipitate of a definite solubility. Taken from von Weimarn, Text Ref. No. 7.

solution, producing a cheesy paste which may be turned upside down. A portion of this on dilution again passes freely through filter paper as a deep blue sol.

Lastly, von Weimarn's diagrams emphasize the importance of extreme insolubility in producing colloidal particles of this unstable type. Even so, the addition of suspending or protective substances (for Prussian blue, oxalic acid; for barium sulfate, potassium citrate) as peptizers hinders the aggregation of the primary particles produced. Barium sulfate is too soluble to retain its colloidal subdivision unless a large amount of alcohol is added to lower its solubility. Unstable sols are usually of extremely low concentration, of the order of 0.1%.

An example of a very dilute unstable or hydrophobic colloid is von Weimarn's sulfur produced by throwing a solution of sulfur in alcohol into a larger body of water in which it is insoluble, the supersaturated solution immediately producing fine particles of sulfur in suspension. These form coarser particles or aggregates on warming, but on heating further they partially dissolve and leave a still finer suspension. On the other hand, Carey Lea prepared concentrated silver sols from silver nitrate by reduction with such stabilizing agents as alkaline dextrin. A whole series of metallic sols has been prepared for medicinal purposes by adding sufficient protective colloid so that the sols may be dried and redissolved in water.

In general, scrupulous attention must be paid to the cleanliness and insolubility of the glass utensils, the purity of the chemicals and the water, and even the atmosphere with which the solutions come in contact. Where no attention is paid to the stabilizing agent, or it is left to chance, the results are usually irregular.

Colloidal particles are often formed from reactions which ordinarily appear insensible, as for example the solution of ordinary silver in

water, whereupon reduction or the action of light produces sufficient colloidal silver to exert a disinfectant action. This has been made use of in sterilizing water by mere contact with metallic silver or plastics containing silver. Colloidal particles are often produced from the electrodes during electrolysis of aqueous solution; Fink [5] states that colloid is also an intermediate stage in electrodeposition from ions in solution.

Colloidal Gold [6]

What the frog was in physiology, colloidal gold was in the early development of the subject of colloids. Colloidal gold is prepared by reduction of gold chloride, preferably pure $HAuCl_4 \cdot 3H_2O$, rather than the gold chloride of photography, $NaAuCl_4 \cdot 2H_2O$. Faraday in 1857 introduced the method of reduction by phosphorus. When 2.5 cc of 0.6% chlorauric acid and 3–3.5 cc of 0.18 N potassium carbonate in 120 cc of water are treated with 0.5 cc of a solution of phosphorus in ether (prepared by diluting the saturated solution to five times its volume), a red sol of gold results whose particles are too small to be seen in the ultramicroscope. This is now known as Zsigmondy's nuclear gold sol. It is of value when larger particles of gold or silver of uniform size are required, for when their salts are reduced in the presence of these previously provided nuclei, all grow alike, and the formation of fresh nuclei is hindered by adding some potassium ferro- or ferricyanide or ammonium salts. Coarser particles are produced when fewer nuclei are added. Various nuclei can be gold- or silver-plated until they grow to any desired size. Citrates and some dyes have the property of accelerating the formation of gold nuclei. On the other hand, practically all additions tend to slow down the rate of crystallization, and hence growth, of existing nuclei.

Many other reducing agents have been employed, some requiring elaborate care and others not, and the colors produced vary from clear ruby red through other reds to brown, orange, purple, and blue. The color of ruby glass is due to very fine colloidal gold. A similar reduction and color is produced in warming gold chloride in glycerin. Reduction at room temperature with hydrazine, hydroquinone, pyrogallol, and gallic acid may give a clear dark blue sol, but sometimes a brown or pink one.

Zsigmondy's red sol [7] is prepared according to elaborate instructions by reduction with formaldehyde but does not always give uniformly reproducible results. Many other recipes have been published.

Hatschek has prepared gold artificially in all the forms in which it occurs in nature.

Nordenson [8] studied the process of reduction by hydrogen peroxide which proceeds according to the equation:

$$2HAuCl_4 + 3H_2O_2 = 2Au + 8HCl + 3O_2$$
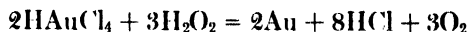
It was followed by measuring the increase of electrical conductivity. When no nuclei were present, the conductivity increased rapidly at first, and then very slowly, until it again suddenly rose to the final value just after the red color appeared. This was explained as due to the supersaturation being released by the appearance of nuclei, a supposition which was confirmed by showing that the reaction went rapidly to completion whenever nuclei were added or were originally present.

It is evident that much of the uncertainty in the preparation of gold sols is dependent upon the degree of hydrolysis of the auric compounds. Moreover, Pauli [9] has found that the stabilizing agents are not auric derivatives, but rather chloraurous acid, $HAuCl_2$, or its salts.

Carey Lea [10] in 1889 attracted much attention with his many preparations of different colors of colloidal silver, one of which "when dry exactly resembles metallic gold." He reduced silver nitrate with iron citrates or tartrates in concentrated solution. Such older publications are worth referring to for the wealth of factual information that is now largely forgotten. Carey Lea also produced colloidal gold by reduction with thiosulfate.

Condensation of Heated Vapors [11]

The same principles regarding formation and growth of nuclei appear in the sudden condensation of supersaturated vapors. If certain organic substances containing platinum or gold are heated in a test tube they decompose, and the cooler glass is covered with a bright platinum or gold mirror.[12]

It should be pointed out that almost all metallic films and mirrors consist of colloidal particles lying side by side but not necessarily touching. Many references are found in the paper by Richter.[13]

Bredig in 1898 introduced the method of forming a supersaturated metallic vapor by submerging an arc passing between two wires of platinum, or other metal, in water, or alcohol, or other liquid. He used direct current of 2-10 amperes. The arc sends out clouds of brown particles of crystalline platinum, the finest of which come from the

metallic gas, with much coarser ones from the molten crater of the arc. Pennycuick,[14] by analyzing the frozen solid, obtained with a composition of the Bredig particles (See Chap. 1, Fig. 1·2) 1 mole of $H_2Pt(OH)_6$ to 3.9 atoms Pt and 1.16 moles of PtO_2. The acid combined in the micelles was 22.5% of its weight, whereas 4.2% was free in solution.

The nobler and less volatile the metal, the easier and finer the preparation, but even sodium can be used in dry ether or pentane at $-100°$ in a hydrogen atmosphere. Many variants of the procedure and electrical equipment have been adopted.[15] The only thorough method of protecting the liquid from direct contact with the arc without introducing still further impurities was that of Giles and Salmon,[16] who used a vertical hollow cross of the same metal, two horizontal arms of which held the wires of the arc while the superheated vapors from it were blown down into the liquid along the other two. In some cases they added a small amount of rubber to stabilize the particles in alcohol.

Again, traces of impurity and of stabilizing agents, even a trace of alkali, are of great importance. However, at least one paper [17] emphasizes that colloids of the unstable type may be successfully prepared in the presence of concentrated electrolytes. Indeed, clay is peptized by concentrated magnesium chloride and the similar action of chloride is the basis of one of the most important recent advances in the technology of drilling muds.

Stable Colloidal Sulfur

Von Weimarn's unstable bluish-white sulfur sol already referred to is readily and irreversibly precipitated by traces of salts. In contrast, Sven Oden has devoted an extensive study [18] to the stable sulfur sols prepared by Selmi in 1844 and Raffo in 1908. Selmi's sol is formed according to the equation:

$$2H_2S + SO_2 = 3S + H_2O$$
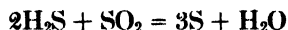
However, Freundlich has pointed out that this is not the sole reaction, but that the stabilizing agent, pentathionic acid, $H_2S_5O_6$, is also formed. Indeed, in very concentrated solutions half the yield may be pentathionic acid, whereupon the sulfur appears in very fine, very stable particles. In dilute solution much of the sulfur is precipitated and the remainder is not very stable. Raffo's sol is prepared by adding concentrated sulfuric acid to a cold, concentrated solution of sodium thiosulfate, where again pentathionic acid is formed. Both sols are purified

by precipitation with a strong solution of sodium chloride, removing the liquid from the sulfur by a centrifuge, filter, or siphon, and redissolving it in hot water.

La Mer uses 0.003 M hydrochloric acid with 0.001 M sodium thiosulfate to produce monodisperse sulfur. (Cf. Chap. 6.) This remains homogeneous and clear for 69 minutes then turns blue for one hour, then successively grey, red, green, etc., as the particles grow up to 700 Å in diameter, and finally beyond.

Hydrolysis

Most solutions of salts of heavy metals contain colloidal particles produced by hydrolysis. The hydrolysis is promoted by boiling or by allowing the crystalloidal materials to diffuse away through a parchment or other similar semipermeable colloidal membrane sufficiently dense to hold back the colloidal particles only (dialysis). This is the usual method of preparing colloidal ferric hydroxide. However, the preparation fails unless an appreciable proportion of ferric salt, such as oxychloride, is retained as stabilizing agent.

Dialysis, or better, electrodialysis (see Chap. 8), is the usual method of purifying colloids by removing crystalloidal impurities.[19] A good means of accomplishing dialysis without diluting the sample is to fill with a solution a length of artificial sausage skin, such as Visking, tie both ends, and place it in water. It is an advantage to place a glass ball in the solution and use a gentle rocking to keep it stirred. The water must be kept pure by replacing it or keeping it running very slowly past the membrane.

Another way of removing electrolytic impurities is by the use of ion exchange resins (Chap. 3). Careful flushing may also serve. Forsyth and Fraser [20] destroyed the colloid gels of soil by freezing the acid precipitated flocculae of humic acids and then thawing. The materials lose their capacity for holding water, and the impurities thus set free can be washed away. Electrolytes are necessary for the freezing effect to take place, since in their absence the material reverts to its colloidal form on thawing.

Peptization

Peptization is the term used by Thomas Graham to describe the process of separating from each other and suspending in a liquid particles of a colloid which have aggregated to form a gelatinous or finely

granular mass. This is of importance in the ceramic industry where
the particles of clay are peptized by aqueous sodium carbonate and
water glass. Graham found that only one part of sodium hydroxide
in 10,000 parts of water will peptize 200 parts of gelatinized silicic acid
in one hour at 100°. Peptization is unpleasantly familiar in the process
of washing out precipitates until they pass through the filter paper,
either by removal of coagulating salts, or, as in the production of sols
of cadmium sulfide, by washing the precipitate with sufficient solution
of peptizing hydrogen sulfide to form a liquid which is pale golden
yellow by transmitted light and greenish by reflected light.

Other Reactions

One colloidal sol may be formed from another, as when chlorine
reacts with the particles of a gold sol or when two colloids are mixed,
as discussed in Chapter 12. Colloidal particles may even be passed
from one solvent into another, as in the manufacture of "Oildag"
from "Aquadag," Acheson's aqueous suspension of graphite stabilized
by vegetable extract. An interesting lecture experiment is to pass a
stream of fine water droplets through a smoky flame. The liquid col-
lected is a clear dark brown sol of carbon particles, which quickly
coagulate when no special stabilizing agent has been added.

An important commercial problem is the preparation of a suspen-
sion of fine coal particles in a petroleum oil for use in oil burners. Most
attempts have been made with heavy-metal soaps as stabilizing agents.
Perhaps the most successful practical solution has been to use high
boiling aromatic solvents from brown coal tar in a naphthene base
heavy oil, resulting in a permanent suspension. This may contain as
much as 30% of carbon in particles between 0.1 and 10 microns.

Milk glass is the suspension of calcium fluoride in glass. Genuine
or synthetic ruby contains chromic oxide in alumina; sapphire, cobalt
oxide in alumina; blue rock salt is colloidal sodium in sodium chloride.

Heyrovsky has developed a significant lecture experiment in which,
when eight volts is applied between mercury electrodes in an aqueous
solution of tetraethylammonium hydroxide (pure), the surface tension
is so reduced, or even made negative, that the mercury streams into
the water as a brown cloud of colloidal particles.

Colloids may be formed by employing a machine similar to a colloid
mill which uses a viscous medium such as dextrin to shear the sub-
stances into colloidal subdivision. The dextrin is then dissolved away
and the suspension transferred to any desired medium by preferential
wetting.

The preparations described above are all for sols. Colloids occur also as jellies, curds, precipitates, coagula, films, fibers, membranes, and other solid structures, and of all these forms sols are often the least important from a practical standpoint. This is evident in such industries as ceramics, textiles, paper, cement, leather, or rubber. The structure of jellies is discussed in Chapter 11, and the special properties of some of these other materials in Chapters 21 to 25.

REFERENCES

The Preparation of Colloidal Sols

SVEDBERG, THE., *Die Methoden zur Herstellung Kolloiden Lösungen Anorganischer Stoffe*, 2nd ed., T. Steinkopff, Dresden and Leipzig (1920)

WEISER, H. B., *Inorganic Colloid Chemistry*, John Wiley & Sons, Inc., New York (1938); Chapman and Hall, Ltd., London. Vols. I, II, and III, *The Colloidal Elements* (1933); *The Hydrous Oxides and Hydroxides* (1935); *The Colloidal Salts* (1938)

Text References

1. GILES, J. K. and SALMON, C. S., *J. Chem. Soc.*, **123**, 1597 (1923)
2. WALDEN, P., *Kolloid-Z.*, **27**, 97 (1920)
3. KRUELEN, D. J. W., *Grundzüge der Chemie und Systematik der Kohlen*, D. B. Centen's Uitgevers-Maatschappij, N. V., Amsterdam (1935); GILLET, A., *Brennstoff-Chemie*, **17**, 421 (1936)
4. LACHS, H. and GESTEL, K., *Z. physik. Chem.*, Abt. A., **137**, 193 (1928)
5. FINK, C. G., *J. Phys. Chem.*, **46**, 76 (1942)
6. WEISER, *op. cit.*
   ZSIGMONDY-THIESSEN, *Das Kolloide Gold*, Akad. Verlag, G.m.b.H., Leipzig (1925)
7. ZSIGMONDY, R., *Lieb. Ann.*, **301**, 29 (1898); *Z. anal. Chem.*, **40**, 712 (1901); cf. VON WEIMARN, *Die Allgemeinheit des Kolloidzustandes*, Theodor Steinkopff, Dresden and Leipzig (1925), Band I, p. 216
8. NORDENSON, H., *Ueber die Bedeutung des Lichts für die Bildung und Stabilität Kolloider Lösungen*, Diss., Upsala (1914)
9. PAULI, W., *Naturwissenschaften*, **20**, 551 (1932), also cf., *Ind. Eng. Chem.* (News Ed.), **10**, 162 (1932)
10. LEA, C., *Am. J. Sci.*, **37**, 476 (1889)
11. SVEDBERG, THE, *Colloid Chemistry*, Chemical Catalog Co., New York (1928)
12. For example, GIBSON, C. S., *Nature*, **140**, 583 (1937); MENZIES, R. C., *J. Chem. Soc.*, 565 (1928), used trimethyl platinum acetylacetone.
13. RICHTER, K., *Kolloid-Z.*, **61**, 208 (1932)
14. PENNYCUICK, S. W., *J. Am. Chem. Soc.*, **61**, 2234 (1939)
15. Compare the account in SVEDBERG, T., *Colloid Chemistry*, *op. cit.* pp. 39–52
16. GILES and SALMON, *op. cit.*
17. VOET, A., *J. Phys. Chem.*, **40**, 307 (1936)
18. ODEN, S., *Der Kolloide Schwefel*, Akademische Buchhandlung, Upsala (1913)
19. STAMBERGER, P., *J. Coll. Sci.*, **1**, 93 (1946)
20. FORSYTH, W. G. C. and FRASER, G. K., *Nature*, **160**, 607 (1947)

CHAPTER **6**

OPTICAL PROPERTIES AND STUDY

OF COLLOIDS

Among striking properties that are developed to a maximum extent by matter in the colloidal condition, one of the most significant has been opalescence. We have already noted the intensity, beauty, and variety of colors which may be manifested by a single material, such as colloidal gold. A solution of one of its salts, so dilute as to be practically colorless, on reduction gives deeply colored colloidal gold, which according to the precise conditions of the experiment may be ruby red, pink, orange, green, yellow, dark brown, clear blue, purple, or other intermediate shades. As will be seen, these colors are again very different by reflected and transmitted light or as seen in the ultramicroscope.

Figure 6·1 indicates diagrammatically after Wolfgang Ostwald the fact that certain properties, such as opacity, increase steadily with increase of particle size from that of a molecule up to visible dimensions. Others, such as Brownian movement, increase steadily in the opposite direction, that is, with fineness of particle size. Still others are at a maximum in the region of colloidal dimensions.

Opalescence

The opalescence of gold sols was noted by Faraday. He rightly contended, with many after him, such as Bredig, that this revealed the particulate nature or heterogeneity of colloidal sols. The opalescence is particularly prominent when a narrow beam of light passes through a colloidal solution, like a ray of sunlight revealing the dancing motes of dust in a darkened room. This is commonly referred to as the

Tyndall effect or cone. Tyndall noticed that all such opalescent light is partially polarized, particularly when viewed perpendicularly to the path of the original ray. A Nicol prism or a polarizing sheet (Polaroid) may be used for seeing through, or in photography for cutting out, atmospheric haze, as in a distant landscape.
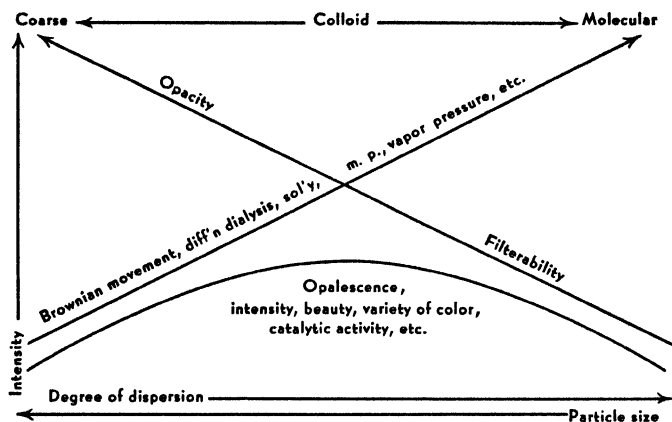
FIGURE 6·1 Diagram showing nature of dependence of physical properties upon particle size, from that of molecules to visible dimensions.

Formerly it was thought that opalescence was a property only of particles, but now it is clearly recognized that light is scattered by all molecules, even in a gas. This scattering of light, or turbidity, is opalescence. The fact that scattered light is polarized is referred to as the Tyndall effect. The following table may serve to emphasize the difference between opalescence and two other somewhat similar optical effects in which we are not here interested:

Opalescence = { Strong
Polarized
Same wave length as incident light

Fluorescence = { Strong
Not polarized
Wave length independent of that of incident light

Raman effect = { Extremely feeble
Polarized
Lines and bands whose wave length changes with incident light

It has been known since the 19th century,[1] and recently rediscovered,[2] that opalescence occurs over a wide range of temperatures above

the critical point of gases and vapors and above the critical temperature of complete miscibility of liquids. Such opalescence is due to evanescent association to form particles of colloidal magnitude.

The opalescence of aqueous solutions of salts of the heavy metals shows the presence of colloidal products due to hydrolysis, which are often suppressed by sufficient addition of acid.

Measurement of Opalescence and Turbidity

The intensity of light scattered by colloidal particles may be measured by several different instruments.

The Tyndall meter [3] measures the intensity of light scattered from the light beams passing through a solution. This is found to be directly proportional to the concentration of the dispersed phase. However, particle size as well as number of particles affects Tyndall values.[4]

A *nephelometer*, such as the commercial Zeiss-Pulfrich Stuphophotometer, is designed for samples of slight opalescence. An intense light passes through the solution at right angles to the line of observation and direct comparison is made with a solution of a standard suspension.

A *turbidity meter* is designed for more opalescent colloids. In the usual instruments direct vision is used, and the light passing from below through the unknown and through the standard in parallel is compared. Another excellent design is that patented by K. B. Barnes of the American Cyanamid Corporation Research Laboratories, Stamford, Connecticut.[5] Here the turbidity of the sample is obtained by the ratio of the total light transmitted to the amount of light scattered, the measurements being made with photoelectric cells. The advantage of this instrument is that it is equally useful for colorless as well as colored light absorbing solutions.

In connection with the recent refinements of the interpretation of light scattering, measurements are now made of the absolute intensity of the scattered light by photo tubes placed at different angles, such as 90°, 180°, 50°, and 130° to the direction of incident light.[6]

Likewise Debye has devised a very sensitive differential refractometer for measuring the difference in refractive index between a colloidal solution and its solvent.

Turbidity measurements have been applied recently [7] to the determination of the molecular weight of micelles of the surface active agent Aerosol MA (*p*-diphenyl sodium sulfosuccinate) using the turbidometric method originated by Debye.[8]

Lord Rayleigh [9] deduced the following formula for the intensity of light scattered by completely independent spherical particles in a

medium which also is an electrical insulator, at a given distance from
the illuminated particles

$$I = kvV^2/\lambda^4$$

where v is the number of particles, V the volume of one particle, and
λ the wave length of the incident light within the medium. The total
scattered light I is related to the original intensity of illumination I_o
by the following formula,

$$\frac{I}{I_o} = \frac{24\pi^3 v V^2}{4} - \left(\frac{n_1^2 - n^2}{n_1^2 + 2n^2}\right)$$

where n is the refractive index of the medium and n_1 that of the particle.

Scrutiny of this formula shows that the shortest wave lengths are
scattered most, for example, violet, ultraviolet, and x rays, as compared
with red light. The most perfect example of scattering of blue light is
the blue sky, which otherwise would be black. This contrasts sharply
with a red sunset which is viewed toward the sun. Another is blue sea
water (noted by Raman), which loses in transparency through the
presence of clouds of organic particles what it gains in their colored
scattered light.

It should be strongly emphasized that often a highly opalescent,
turbid solution appears as clear as water when one looks directly
through it toward the source of transmitted light, such as a window.
Other examples of scattering are the bluish white of watered milk or
of gum gamboge or sulfur sols. Mecklenburg first used the Rayleigh
formula to determine the size of such spherical nonconducting sulfur
particles. Tezak [10] describes the simple relations between absorption
and scattering for this simplest case.

Great departures [11] from the above simplified formula and from
the predicted color are observed where conductors are present, as in
most sols, or where the particles are not spheres, or where they are
aggregated, or where absorbing materials are present. Secondary
scattering and diffraction effects produce further complications.
Polarization is often elliptical, or even circular. Krishnan [12] has studied
the depolarization of Tyndall scattering; Smirnow and Mecklenburg [13]
have shown that even in the simplest cases the fourth power law holds
only for particles less than 500 Å or sometimes 1000 Å in diameter.
Above that, the exponent very rapidly diminishes. Ray (1920) and
La Mer and Johnson [14] (1945) found that the light scattered from a
monodisperse sulfur sol can exhibit different colors when viewed at
different angles to the incident beam. La Mer and Barnes [15] have

prepared sulfur sols of very uniform size which change from pale blue in all directions to higher orders of the spectrum with increase in particle size, of which colors a clear, intense red is the most striking.

> The purity of the spectral colors to the naked eye furnishes a trained observer with a ready qualitative estimate of the degree of monodispersity, while the number of spectral orders gives a ready rough measure of the size. In the size range in which the orders occur, brilliant colors imply a strictly monodisperse character, while pastel shades indicate less monodisperse distributions. Opalescence is a mark of polydispersity.[16]

Further Development in the Theory of Light Scattering

Debye [17] has summarized the present state of knowledge as to how light scattering can assist in the study of the size and shape of colloidal particles and describes numerous uses for this property, including determination of molecular weights of high polymers in solution. It is necessary to measure the ratio of the scattered intensity to the incident intensity at a given scattering angle and the asymmetry of the scattered radiation about the perpendicular to the incident beam. It is usually necessary to make measurements at several concentrations of the dispersed phase and extrapolate to infinite dilution. It is also necessary to measure the rate of increase of the refractive index of the solution or suspension with increasing concentration of dispersed phase.

An exceptionally simple, comprehensive, and lucid account of light scattering of solutions and sols has been given by Mark.[18] A more general and recent review is that of Oster.[19]

The compelling interest in the quantitative characterization of high polymers, particularly in extreme dilution, led to a reexamination and refinement of the light scattering equations of Rayleigh, Einstein, and Debye.

The following factors are taken into account: the interactions between molecules in a condensed system which are no longer isolated independent scattering centers; the scattering by molecules, none of whose dimensions approaches that of a wave length of light λ, as contrasted with particles one or more of whose dimensions is comparable with λ. The result is different for rod-shaped, spherical, and for long, coiled molecules embedded in solvent. In all these cases the effect of concentration and, in particular, the relative effect of various angles of observation relative to the incident beam of light, are studied and their significance deduced. Further refinements of calculation and interpretation are in progress.

From these observations and deductions we may quote a few salient results and formulas.

In the first place, it is pointed out that if one could have an ideal, colorless crystal in which the molecules were strictly aligned in a space lattice without any Brownian movement, there would be no light scattering at all. Second, scattering from a gas or pure liquid is very small as compared with even the most transparent of colloidal solutions. Hence ordinary scattering relates chiefly to the colloidal particles. This is evident from the following table taken from Mark,[20] in which D is the length of a column of liquid in which, through scattering, the intensity of the transmitted light has been reduced to $1/e$ of its initial value.

TABLE 5

ORDER OF MAGNITUDE OF LIGHT SCATTERING IN A FEW TYPICAL SYSTEMS

System	Order of magnitude of turbidity, cm^{-1}	Order of magnitude of D, various units
Ideal crystal	Virtually zero	Infinite
Ideal gas or vapor	10^{-7}	100 km
Pure liquids	10^{-6}	1 km
Solutes of ordinary small molecules in 1% concentration	10^{-4}	100 m
Polymer solutions in 1% concentration	10^{-3}	10 m
Rubber latex in 1% concentration	10^{-1}	10 cm
Milk	10	1 mm

Scattering is the result of fluctuations in density or in concentration due to Brownian motion of the molecules of the particles and even of different segments of a chainlike particle. The unit of thermal energy kT may be compared with the work required to accomplish a corresponding density change by outside pressure or with a change in refractive index by compression, or it may be compared with the osmotic work required to produce a change in concentration. Thus these fluctuations which are revealed by light scattering are quantitatively linked to osmotic pressure and, hence, molecular weight.

However, it should be remembered that the equations presently in use for high polymers, based on fluctuations, are restricted to particles very small in comparison with the wave length of light.

In determining the molecular weight M from osmotic pressure π, the measured values are inserted in the following linear equation, where c is the concentration and B a constant and the value of M is taken as that at infinite dilution:

$$\frac{\pi}{c} = RT/M + Bc$$

Correspondingly for light scattering

$$Hc/\tau = \frac{1}{M} + \frac{2Bc}{RT}$$

The term on the left is the reduced transparency, H being a factor depending upon wave length and refractive index; τ is the total fraction of incident light scattered per unit volume. The results for these two equations, after correcting for a variety of influences, are in satisfactory agreement. Results are also obtained for the shapes and flexibility of the primary colloidal particles.

Light scattering gives a weight average molecular weight, because the larger particles scatter more. On the other hand, osmotic pressure gives a number average, because every ion, molecule, or particle gives the same osmotic effect. The molecular weight of egg albumin determined by light scattering is 34,500.[21]

The shape of a polymer molecule is found to depend upon how similar it is to the solvent. For example, polystyrene in toluene [22] is fully solvated, but in methyl ethyl ketone is coiled up so that different parts of the same molecule may stick together — associate — and, presumably, when different molecules touch, they too must associate. The temperature dependence of the results for polyvinyl chloride demonstrates a marked tendency toward aggregation and association.[23]
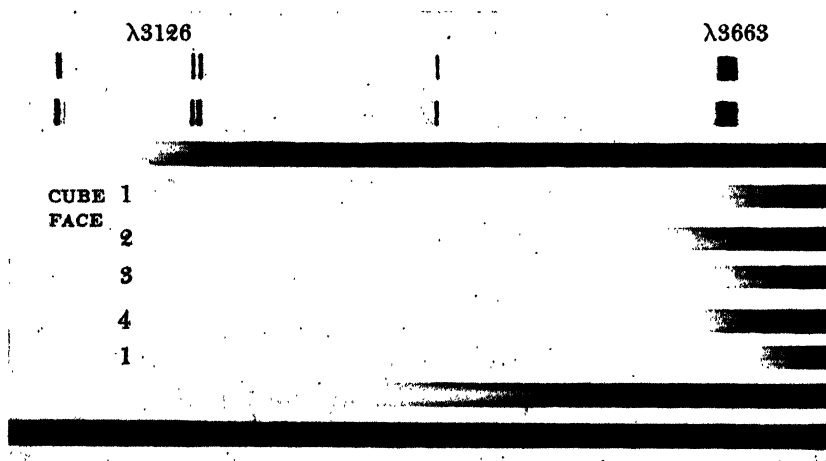
FIGURE 6-2 A single concentration of sodium oleate (0.370 N) within a quartz cube, photographed using the hydrogen-discharge tube, each picture being taken at right angles to the one preceding it, water first and last as reference.

Light absorption, as well as light scattering, also takes place in colloid sols. It was found that the absorption of light by sodium oleate solution depended on orientation of the colloidal particles. Micelles oriented in one direction absorbed more ultraviolet light than the same micelles turned at right angles, as is strikingly illustrated by Figure 6·2.[24]

Examination of Shape with Polarized Light

Certain colloidal solutions become doubly refracting when they are stirred or allowed to stream.[25] On examination between crossed Nicols or crossed Polaroids (see later), which normally extinguish each other completely, the field lights up and the stream lines appear on account of the scattering and repolarization of the incident polarized light. Such sols must contain particles that are oriented by flow. Hence they cannot be spherical but must be either "rodlike" or "disc-shaped." A good specimen for examination is vanadium pentoxide sol which has aged by standing for a year or more or by boiling for an hour. Rods will place themselves parallel with the stream lines, but plates will place themselves at right angles to them, except near the wall of a tube, to which they will tend to be parallel with the longest axis in the direction of flow.

The streaming sol may be illuminated by polarized light from any one of three directions with the plane of polarization either horizontal or vertical. It is evident that the scattering from the end and sides of a rod or from the edges or the flat face of a disc will be distinguished by corresponding differences in intensity. These differences in amount of reflection, especially of polarized light and in colored sols, may produce dichroism even when the particles themselves are neither dichroic nor anisotropic.

Thus it has been found again that most sulfur sols are "spheres" or granules. So are certain sols of arsenic trisulfide and fine particles of gold and silver. Vanadium pentoxide, aniline blue, and benzopurpurin are rod shaped, while ferric hydroxide and blue gold form discs or leaves. The latter may be immensely larger than the primary particles of which they are composed. Suspensions of stearic acid have rodlike particles when the acid is solid, but when the suspension is heated above the melting point of the acid, the particles become spherical droplets and do not show streaming double refraction until cooled again.[26]

An interesting example of rod-shaped particles is the Stanley protein which constitutes the tobacco mosaic virus, as shown by the

stream double refraction of suspensions of the virus. Takahashi and Rawlins [27] used the polarizing microscope shown in Figure 6·3 with crossed polarizer P and analyzer A placed at an angle of 45° to the stream flowing into or out of the pipette B; in the former case, suction, a dark cross would be produced only by rodlike particles while neither double refraction nor a dark cross would be produced by discs.

Krishnan used horizontally polarized incident light and then compared the ratios of the vertical to the horizontal components in the scattered light. He has derived a theory [28] connecting the result with the size of the scattering units as well as their deviation from sphericity. Krishnan's theory is succinctly abstracted by Gehman and Field.[29] Krishnan [30] also found evidence of molecular aggregation in glasses by the reversal of polarization produced when the incident beam is linearly polarized in the direction of the scattering plane.

Björnstahl in 1921 used a magnetic field for testing for double refraction and again found that Oden's sulfur gave none and therefore was spherical. Sulfur sols produced by grinding or by von Weimarn's method proved to be asymmetrical. He found that all gold particles over 100 Å in diameter were not symmetrical. König [31]

FIGURE 6·3 Polarizing microscope, diagrammatically sketched. Taken from Takahashi and Rawlins, *Science*, **77**, 26 (1933).

found that ferric hydroxide sol was likewise doubly refracting, although he could not observe it with the streamlining method of Maxwell. Synthetic polystyrenes and cellulose derivatives also show but little double refraction. Particle size can sometimes be determined by magnetic double refraction.[32]

The foregoing phenomena suggest methods of following the rate of formation and growth and of coagulation of colloidal particles. Opalescence is the basis of nephelometry, long used in quantitative chemical analysis.[33] Several other phenomena of colloidal particles and polarization of light cannot be discussed here for lack of space. They include Zocher's experiments, the Weigert effect,[34] the inter-

ferometric studies of light scattering by Venkateswaran,[35] and the spectro-interferometry of Bai.[36]

Polaroid Films

Polaroid films, invented by Land,[37] are an excellent example of elongated particles oriented by streaming in a colloidal medium, which is then solidified. Good descriptions with references are given by Freundlich [38] and by Levey.[39]

Polaroid films have consisted of herapathite, which consists of quinine, hydriodic acid, and sulfuric acid ($4Qu, 3H_2SO_4, 2HI, I_4, 6H_2O$), embedded in cellulose ester pulled or stretched or extruded as a plastic film prior to the removal of volatile solvents. Alternatively, the crystals may be oriented on glass plates. The degree of polarization is up to 98% for light of wave lengths between 5000 and 7000 Å but is less at the violet end where, however, absorption of light is predominant. At the red end the transparency becomes so great that crossed polaroid films make a good filter for transmitting infrared only. More recently, colorless Polaroid films have been made by stretching polyvinyl films during their polymerization.

Another type of polarized film can be made from noncrystalline coatings such as those known as Polocoat. They consist of mixtures of dichroic dyes. The solute is flowed onto a surface and allowed to drain and dry. During drying the molecules tend to fall spontaneously into parallelism with respect to one axis. The degree of polarization is varied by varying the thickness of the applied film, and its configuration is altered by rubbing the surface on which it is poured.

Dielectric Constant of Colloidal Systems

This closely related field has been summarized by Heymann,[40] dealing with nonpolar and polar particles, particle weight, jellies, gels, solids, and boundary layers.

X-Ray Examination

Debye and Scherrer were the first to examine the dried residues from gold sols, applying their powder method and showing that the main part of each particle is a small crystal of gold. Much later, Milligan and Weiser examined sols, as well as an extensive series of gels, finding that in many cases particles give identifiable crystal patterns.

Further, with some that appeared amorphous to x rays, definite diffraction patterns were given by electron beams. Some, however, appear amorphous like glasses.

The Ultramicroscope

The ultramicroscope is an ordinary microscope used to examine opalescence. Colloidal particles, less than a wave length of light in diameter, cannot reflect light but do diffract or scatter it. Thus each particle becomes a center of light which, if sufficiently intense, may be seen in the ultramicroscope, much as a star may be seen by the naked eye although its disc or diameter is far too small to be seen in the most powerful telescope.

The theory of the resolving power of the microscope due to Abbe leads to the conclusion that one cannot resolve or distinguish points closer than $\lambda/2n \sin \alpha$, where λ is the wave length of the light, n is the refractive index of the medium in contact with the lens on the side nearer the object, and α is one half the angle at which the diameter of the lens is seen if viewed from the particle. The term $n \sin \alpha$ is called "numerical aperture." Hence, to get a true image of the smallest microscopic particle, the numerical aperture must be made as large as possible. Sin α cannot exceed the value 1, in practice, say 0.95. Therefore, the lens is immersed in a liquid to increase n. Water, whose refractive index is 1.33, yields a practical value of 1.25, and cedar oil, 1.4. The last recourse is to ultraviolet light and photography, as in the microscope of J. E. Barnard of London or F. F. Lucas of the Bell Telephone Company, the final limit being 1100 Å as compared with a minimum 2000 Å for visual examination.

Any arrangement will serve in which no direct light enters the microscope but only that diffracted from the particles (dark ground illumination). The visibility of the particles depends upon the specific intensity of illumination, the best being sunlight, though an arc or point light source is usually employed. It depends on the contrast with the dark background, and, hence, only the thinnest layer of particles should be illuminated. It depends upon the product of the squares of the numerical apertures of the illuminating and viewing systems, and, finally, as was likewise seen in the foregoing section, it depends upon the difference in refractive index of particles and medium. It is disappointing that many of the most important colloidal sols, and all jellies, even though prepared with utmost care, cannot be resolved into particles with the ultramicroscope. Examples of this are afforded
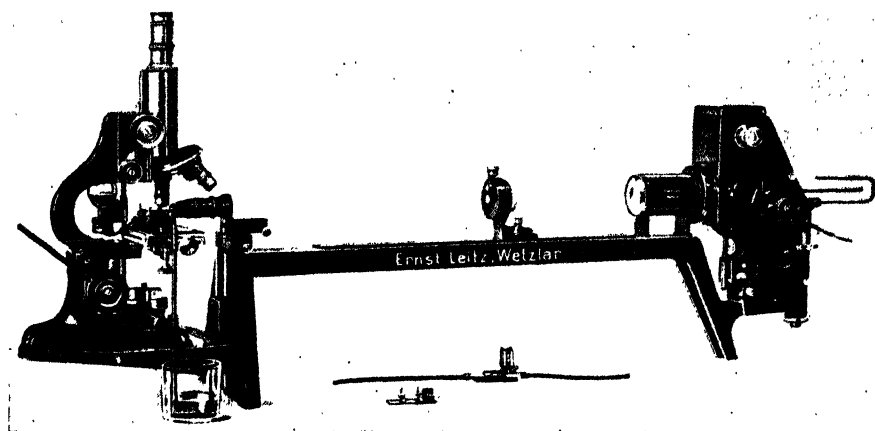
FIGURE 6-4 Siedentopf-Zsigmondy ultramicroscope as supplied by E. Leitz, New York, reproduced with permission from their pamphlet No. 7043.

by soap, ferric hydroxide, magnesium arsenate, and nitrocellulose.
• This failure is attributed to the very small size of the primary particles
and to the modification of their refractive index by solvation.

Of the many forms of ultramicroscope, each of which has its own
advantages and disadvantages, only a few of the current, commercially
available forms need be mentioned. The most diagrammatic is that
supplied by Leitz, illustrated in Figure 6-4. This resembles the Sieden-
topf-Zsigmondy ultramicroscope in which the microscope is at right
angles to the illuminating beam. This instrument has three distinguish-
ing features. The first is the slit, adjustable by a micrometer screw and
capable of being turned to a vertical or horizontal position. The second
is an objective used to reduce the dimensions of the light beam from
the slit, and the third is a cell with a vertical and horizontal face for
illumination and examination, respectively.

With the slit in the horizontal position only a microscopically thin
layer of liquid is illuminated. Its thickness may be measured in terms
of a scale laid in the ocular of the microscope by turning the slit verti-
cally and thus examining it edgewise. Similarly, with the slit now
horizontal, a small area may be measured out and the number of
particles visible in this known volume counted. The ocular scale may
be calibrated by viewing a standard subdivided millimeter. Hence,
if by chemical analysis or evaporation to dryness the total weight of
colloid per cubic centimeter is known, the average weight of each
particle may be directly calculated. The average size can only be

estimated by some such assumption as that the colloidal particle has the same density as the material in bulk and that it is spherical or cubical.

While this ultramicroscope is convenient for examination of a sol, its disadvantages are that the intensity of illumination is none too great and that it is difficult to use it at other than room temperature.

The Zsigmondy immersion ultramicroscope, as supplied by Winkel-Zeiss of Göttingen, is similar, but the intensity of the light is greatly increased by the use of large numerical apertures with high magnification. In order to bring the two objectives simultaneously into focus on the same point, it was found necessary to cut off a portion of the nose of each objective at an angle of 45°. The intensity of the light from the particle is thus increased about 21-fold. Particles of gold down to about 40 Å become visible. It is excellent for photographic purposes. A hanging drop or soft piece of jelly may be directly examined, or a cell may be employed.
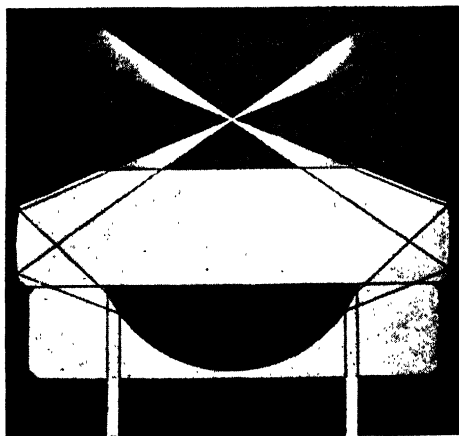
FIGURE 6·5 Diagram indicating the path of light through the dark ground condenser of the cardioid ultramicroscope.

The cardioid ultramicroscope supplied by Carl Zeiss comprises a dark ground condenser in which the path of light is indicated in Figure 6·5. All direct light is cut off by an opaque screen in the condenser. Its advantage is great intensity of light because the focus for truly cardioid surfaces would be a mathematical point and for the actual spherical purposes a very short vertical line. In order to confine the illuminated depth to sufficiently thin layers, the fused silica cell of the form shown

in cross section in Figure 6·6 is employed. Its disadvantage is that the surfaces must be scrupulously clean and free from ultramicroscopic particles, which requires drastic cleansing, completed by heating to redness in a free flame after washing with freshly distilled alcohol. Its great advantage is that if a heating film of platinum is placed on the cover of the cell it may be used at any temperature up to the boiling point of the system.

FIGURE 6·6 Cross section of grooved, fused silica cell (for use with cardioid ultramicroscope), which confines illuminated depth to thin layer.

The photographs of soap curds (gels, not jellies) reproduced in Figure 6·7 were taken by Darke and Salmon [41] in the author's laboratory. A movie camera was found of great service in their work.

The Spierer lens was warmly recommended, especially for solid systems, by Seifriz,[42] and Figures 6·8 and 6·9 are taken from an article by him. Most authorities, however, distrust observations made with it because of confusing extra diffraction lines.

With the cardioid condenser an azimuth diaphragm to restrict the illumination to two sides may be employed. If the moving particles are rodlike or disc-shaped, they appear to scintillate according to the way they light up and come into appropriate positions. A similar effect may be observed with the naked eye [43] with a sol of lead carbonate.

Intensity of illumination is so great with the cardioid and immersion ultramicroscopes that the systems may rapidly undergo photochemical alteration or coagulation.[44]

If particles are invisible in the ultramicroscope, they may sometimes be made visible by gold plating, using Zsigmondy's nuclear method. The chlorauric acid should be less than 1/10,000 N and there should be at least 5×10^9 nuclei per cubic centimeter. The number of particles should be found to be independent of the concentration of chlorauric acid employed. This method has been used even in non-aqueous solutions.

The Electron Microscope

In the preceding section it has been shown that a theoretically perfect optical microscope reaches an absolute limit of resolution at about 2000 Å. The ultramicroscope reveals the presence of smaller
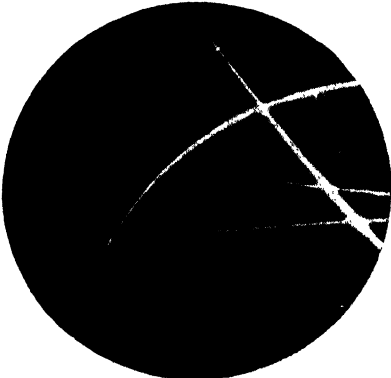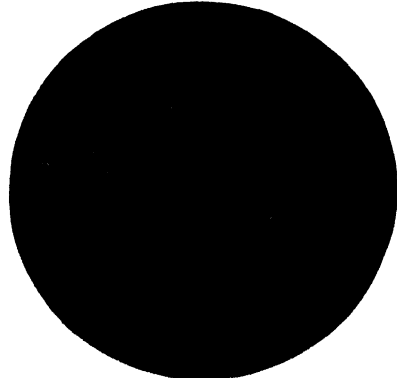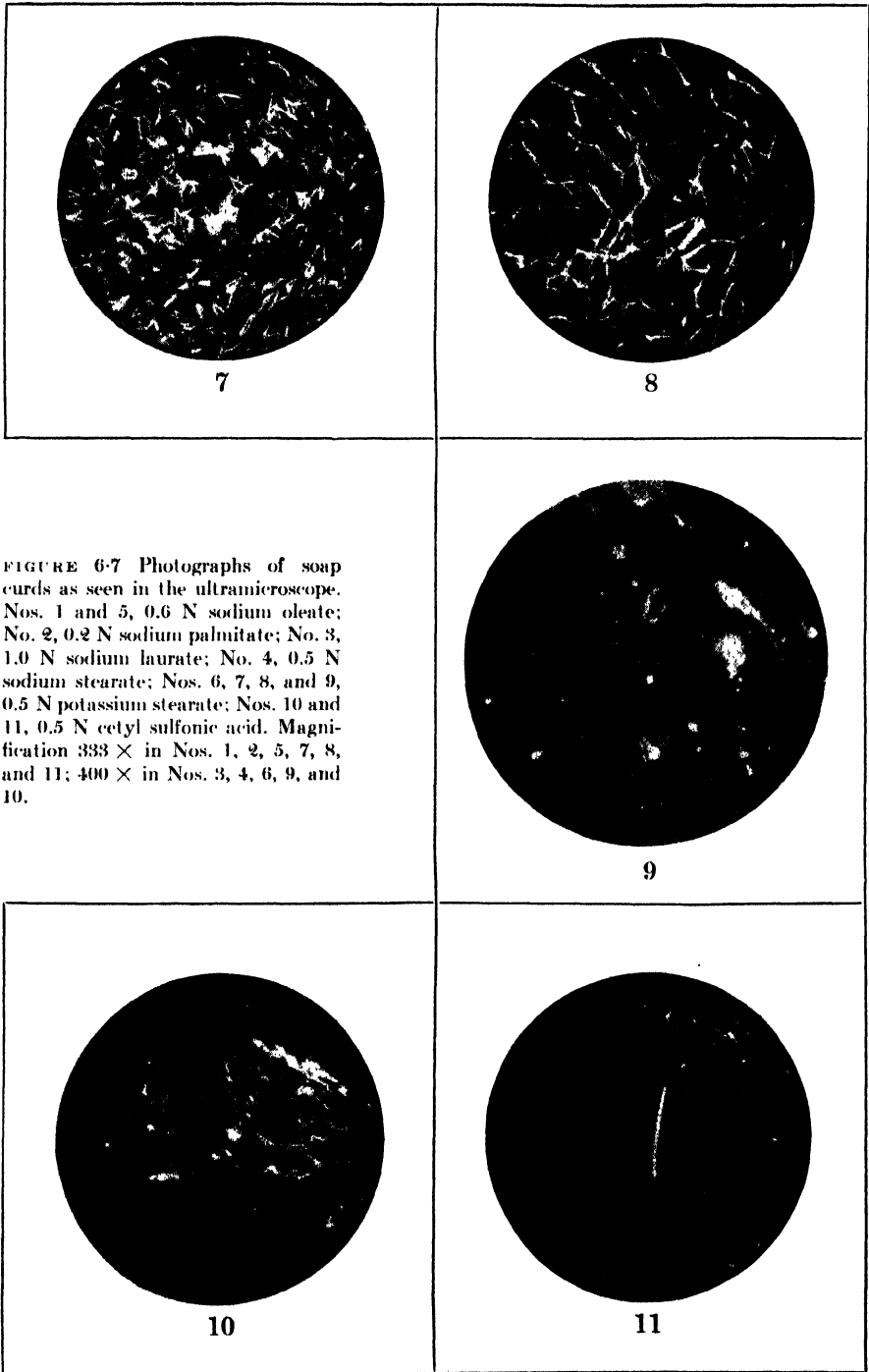
FIGURE 6·7 Photographs of soap curds as seen in the ultramicroscope. Nos. 1 and 5, 0.6 N sodium oleate; No. 2, 0.2 N sodium palmitate; No. 3, 1.0 N sodium laurate; No. 4, 0.5 N sodium stearate; Nos. 6, 7, 8, and 9, 0.5 N potassium stearate; Nos. 10 and 11, 0.5 N cetyl sulfonic acid. Magnification 333 × in Nos. 1, 2, 5, 7, 8, and 11; 400 × in Nos. 3, 4, 6, 9, and 10.

particles, provided they are of refractive index sufficiently different from that of the solvent and the illumination is sufficiently intense — although these conditions are frequently not realized; however, it does not show the size or give any true image of the particle.

FIGURE 6·9 Part of surface of cellulose wall of a living plant (*Allium*) cell taken with Spierer lens, showing perfect parallel orientation of striae, except at extreme left where orientation is disturbed by strain. Taken from Seifriz, Text Ref. No. 42.

FIGURE 6·8 (1) Illustration of the principle of the Spierer lens; (2) ellipse of light diffracted by a colloidal particle. The greatest intensity is in the direction of the illuminating ray (from below).

Electron beams have a wave length incomparably smaller than that of light. Thus with 60,000 electron volts, the wave length is only 0.05 Å or about 100,000 times smaller than ordinary light. Hence *if* a lens with numerical aperture of even 0.01 were possible, one could actually resolve 3 Å. H. Busch in 1927 showed that any cylindrically symmetrical magnetic or electrostatic field acts as a lens toward an electric beam. In 1928 Rupp published the first description, followed by Knoll and Ruska and many others who have constructed electron microscopes. They were obtainable commercially from Siemens and Halske in Germany, and, more recently, have become available from the Radio Corporation of America and other companies. In Germany electron microscopes were entitled Ruṣka's Uebermikroscop. The analogy between optical and electrical microscopes is brought out diagrammatically in Figure 6·10.

Present day electron microscopes operate at potentials between 50,000 and 250,000 electron volts, the latter correspondingly reducing

the wave length of the electrons. They definitely resolve down to 40 Å, with indications down to 10 Å. When and if resolution is improved down to a few Ångström units, all atoms and molecules will be resolved and science will have entered upon a new epoch. The present status of electron microscopy is comparable with that of Galileo's first crude telescope with which the moons of Jupiter were discovered. Some formidable technical difficulties are still to be overcome.

A recent analysis of the underlying theory concludes with the words, "It is therefore to be expected that single atoms will be visible as soon as one of the methods of spherical correction proves to be practically feasible"; [45] and a number of suggestions are put forward to this end.
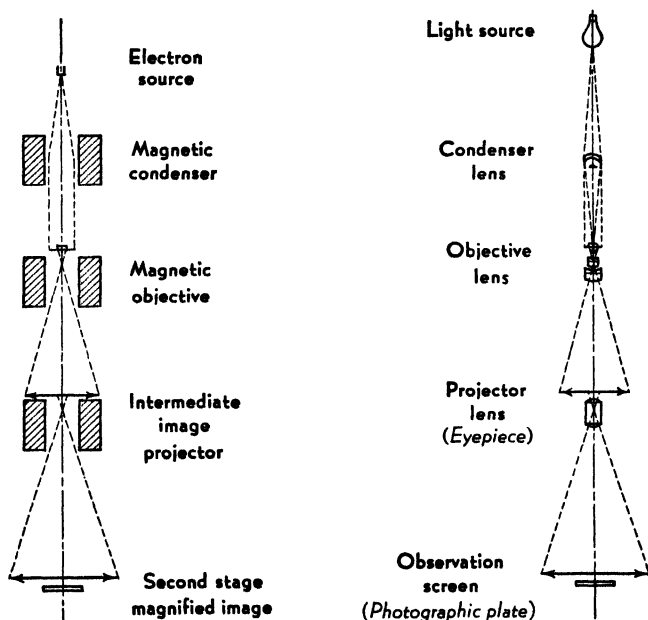
FIGURE 6·10 Diagram showing analogy between electron and optical microscope. Taken from Zworykin, *Science*, 92, 51 (1940).

One severe restriction in the applicability of the electron microscope to the examination of colloid, and, more especially, biological systems is the fact that at present the specimens are almost always exposed to high vacuum, but this is a problem that can surely be solved in several ways. McBain and Abrams [46] have developed closed, liquid-tight cells transparent to electrons for this purpose, but probably dark-field illumination (an "electron ultramicroscope") will be necessary

to distinguish colloidal particles from the medium. Brownian motion is of far too great an amplitude to be readily observed. However, these cells may become indispensable for the examination of sections which can be made with the ultramicrotome 0.1 micron thick, without exposure to vacuum.[47]

FIGURE 6·11 Earliest electron photograph in which ordinary molecules have been partially resolved. Unsupported films of sodium laurate obtained by evaporating a film of sodium laurate solution on fine gauze. The diameters of the fibers are multiples of 40A, the length of pairs of sodium laurate molecules placed end to end at right angles to the fiber.

In Figure 6·11 we present for historical interest the first electron photograph in which ordinary molecules have been partially resolved.[48] Here pairs of sodium laurate molecules, molecular weight 200, lie at right angles to the ribbonlike curd fibers and make their width 40 Å or a multiple thereof. One of the first biological specimens to be photographed is tobacco mosaic virus whose enormous particles are rods with about the proportions of a finger. Remarkable comparisons between photographs of the identical fields taken with the electron
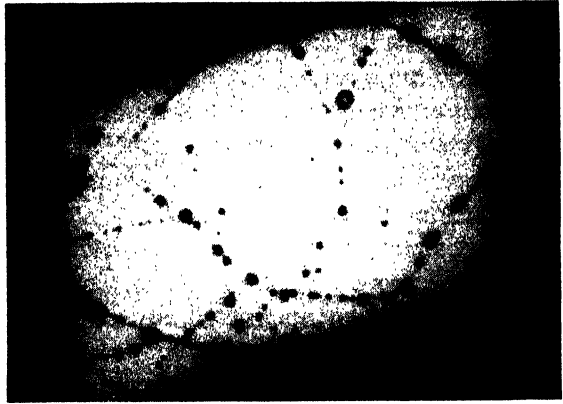
microscope and with the best optical microscope technique have been published by Barnes and Burton.[49]

Hauser and le Beau [50] have shown that some of the electron microscope technique may be adapted to great advantage in examining specimens with intense vertical illumination, in the so-called Ultrapak Lightscope, manufactured by Leitz. If the photograph is taken on Kodachrome film, which is grainless, it may be enlarged very greatly after development on regular printing paper with a light background. One of their photographs of rubber is illustrated in Figure 6·12.

The technique for the examination of the desiccated residue of colloidal systems *in vacuo* is undergoing rapid development. For ex-
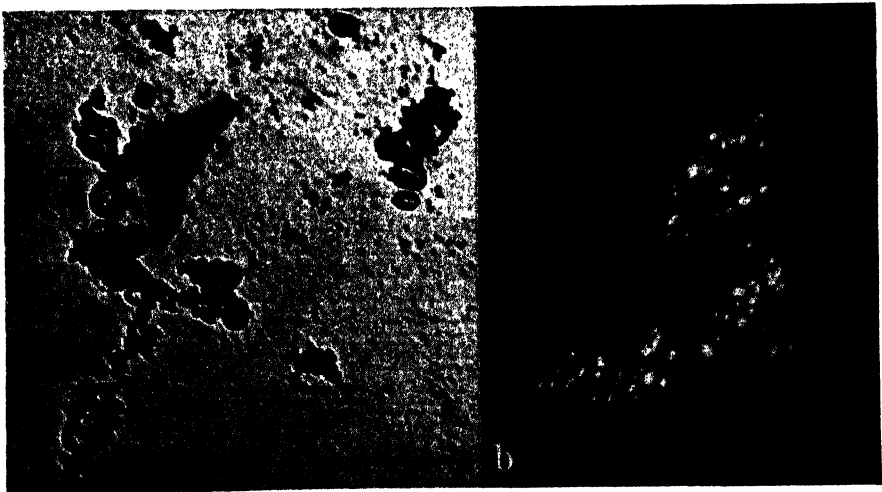
FIGURE 6·13 Electron micrographs of colloidal gold aggregates: (a) bright-field illumination; (b) dark-field illumination. Courtesy *Journal of Applied Physics*. Taken from Hall, C. E., *J. Appl. Phys.*, 19, 198 (1948).

ample, a very effective technique is coating the specimen with gold, chromium, or other metal from one side so as to emphasize contrasts in form. This is called shadow-casting and has been applied to hemo-cyanin particles, biological specimens, and macromolecules.[51] By its means the separate particles or molecules in crystals of tobacco necrosis virus have been resolved and their pseudocubic arrangements and spacings determined.[52]

Dark-field observation has been developed for the electron micro-scope, too. Figure 6·13 compares photographs of dried aggregates of colloidal gold by transmitted electrons and by scattered electrons. In each case the aggregate is of the ramifying aggregate type described in Chapter 10. In the dark-field electron photograph the aggregate as a whole is represented by a diffuse image on which are superimposed sharp, bright diffraction images of discrete crystallites of gold which give a good measure of their shape and size and relative positions.

REFERENCES

Text References

1. References and explanation in F. W. KUSTER and A. THIEL, *Lehrbuch der allgemeiner physikalischen und theoretischen Chemie*, Vol. 1, p. 330–333; KARL WINTER'S *Universität Buchhandlung*, Heidelberg (1913); F. B. YOUNG, *Phil. Mag.*, 20(6), 814 (1910)
2. FRENKEL, I., *Kinetic Theory of Liquid*, Clarendon Press, Oxford (1946), p. 350
3. TOLMAN, R. C. and VLIET, E. B., *J. Am. Chem. Soc.*, 41, 297 (1919)
4. BECHHOLD, H. and HEBLER, F., *Kolloid-Z.*, 31, 70 (1922)
5. STAMBERG, O. E. and MCBAIN, J. W., published by the Tech. Committee of the Master Brewers Association, Oct. 1, 1942
6. DEBYE, P., *J. Appl. Phys.*, 17, 392 (1946)
7. STAMM, R. F., MARINER, T. and DIXON, J. K., *J. Chem. Phys.*, 16, 423 (1948)
8. DEBYE, P., *J. Phys. Coll. Chem.*, 51, 18 (1947)
9. LORD RAYLEIGH, *Phil. Mag.*, 47(5), 337 (1899)
10. TEZAK, B., *Kolloid-Z.*, 74, 16 (1936)
11. MIE, G., *Ann. Physik.*, 25(4), 377 (1908) and DEBYE, P., *ibid.*, 30(4), 57 (1909)
12. KRISHNAN, R. S., *Proc. Indian Acad. Sci.*, 1, 717, 782 (1935); 5, 94, 305, 407, 499, 551, 577 (1937); 7, 21, 91, 98 (1938)
13. SMIRNOW, L. W., *Kolloid-Z.*, 71, 261 (1935); MECKLENBURG, W., *ibid.*, 16, 97 (1915)
14. LA MER, V. K. and JOHNSON, I., *J. Am. Chem. Soc.*, 67, 2055 (1945)
15. LA MER, V. K. and BARNES, M. D., *J. Coll. Sci.*, 1, 71 (1946)
16. LA MER, V. K., *J. Phys. Coll. Chem.*, 52, 65, 68 (1948)
17. DEBYE, P., *J. Phys. Chem.*, 51, 18 (1947)
18. MARK, H., in R. E. BURK and O. GRUMMITT, *Frontiers of Chemistry*, Vol. V, Inter-science Publishers, Inc., New York (1948), pp. 121–173
19. OSTER, G., *Chem. Reviews*, 43, 319–365 (with 227 references) (1948)
20. MARK, H., *Chemical Architecture*, Interscience Publishers, Inc., New York (1948), p. 137
21. PUTZEYS, P. and BROSTEAUX, J., *Trans. Farad. Soc.*, 31, 1314 (1935)
22. DOTY, P. M. and KAUFMAN, H., *J. Phys. Chem.*, 49, 583 (1945)

23. DOTY, P. M., WAGNER, H., and SINGER, S., *J. Phys. Coll. Chem.*, 51, 32 (1947)
24. MCBAIN, M. E. L., *J. Phys. Chem.*, 48, 89 (1944)
25. Although long known and studied (for example, Maxwell 1874), our recent information is due to Freundlich, Zocher, and their collaborators. For systematic discussion and references see H. FREUNDLICH, *Kapillarchemie*, Akademische Verlag, Leipzig (1932) Band II, pp. 40–57. A good discussion is also given in H. AMBRON and A. FREY, *Das Polarisationsmikroskop*, same publisher (1936)
26. RAMAIAH, K. S., *Current Sci.* (India), 8, 360 (1939)
27. TAKAHASHI, W. N. and RAWLINS, T. E., *Proc. Soc. Exp. Biol. Med.*, 30, 155 (1932); *Science*, 77, 26 (1933); *Phytopathology*, 24, 111 (1934)
28. KRISHNAN, R. C., *Proc. Indian Acad. Sci.*, 1, 709, 717, 782 (1935); also KRISHNAN, R. S. and RAO, P. V., *ibid.*, 20, 109 (1944)
29. GEHMAN, S. D. and FIELD, J. E., *Ind. Eng. Chem.*, 29, 796 (1937)
30. KRISHNAN and RAO, *loc. cit.*
31. KÖNIG, W., *Kolloidchem. Beihefte*, 23, 110 (1926); see also SIGNER, R. and CROSS, H., *Z. physik. Chem.*, Abt. A, 165, 161 (1933); 150, 257 (1930)
32. STOTT, F. D. and VON ENZEL, A., *Nature*, 161, 728 (1948)
33. WELLS, P. V., *Chem. Rev.*, 3, 331 (1927)
34. CERDYNCEV, S., *Acta Physicochim.*, 3, 355 (1935)
35. VENKATESWARAN, C. S., *Proc. Indian Acad. Sci.*, 15A, 322 (1946)
36. BAI, K. S., *Proc. Indian Acad. Sci.*, 15A, 338 (1942); 18, 210 (1944)
37. LAND, E. H., British Patent 412,179, 1934
38. FREUNDLICH, H., *Chem. & Ind.*, 56, 698 (1937)
39. LEVEY, H. A., *Ind. Eng. Chem.* (News Ed.), 17, 527 (1939)
40. HEYMANN, E., *Kolloid-Z.*, 66, 229, 358 (1934)
41. DARKE, W. F., MCBAIN, J. W. and SALMON, C. S., *Proc. Roy. Soc.*, 98A, 395 (1921)
42. SEIFRIZ, W., *Ind. Eng. Chem.*, 28, 136 (1936)
43. TRAUBE, I. and KLEIN, P., *Kolloid-Z.*, 30, 19 (1922)
44. HAINES, R. T. M., *Nature*, 139, 32 (1937)
45. SCHEAZER, O., *J. Appl. Physics*, 20, 20 (1949)
46. ABRAMS, I. M. and MCBAIN, J. W., *J. Appl. Phys.*, 15, 607 (1944); *Science*, 100, 273 (1944)
47. O'BRIEN, H. C. and MCBAIN, J. W., *J. Am. Chem. Soc.*, 68, 1139 (1946)
48. MARTON, L., MCBAIN, J. W. and VOLD, R. D., *J. Am. Chem. Soc.*, 63, 1990 (1941)
49. BARNES, R. B. and BURTON, C. J., *Ind. Eng. Chem.*, 35, 120 (1943)
50. HAUSER, E. A. and LE BEAU, D. S., *Ind. Eng. Chem.*, 37, 786 (1945); *J. Am. Chem. Soc.*, 68, 153, 1137 (1946)
51. WILLIAMS, R. C., and WYCKOFF, R. W. G., *J. Appl. Phys.*, 17, 23 (1946)
52. MARKHAM, R., SMITH, K. M., and WYCKOFF, R. W. G., *Nature*, 161, 760 (1948)

BROWNIAN MOVEMENT

When one makes an observation with the ultramicroscope, one is first of all struck with the ceaseless, irregular movement of the particles. Such movement was first noticed by the English botanist Brown in 1827 with pollen grains. He rightly considered it to be a property of all the smallest microscopically visible particles. Throughout the nineteenth century this view was looked upon with great skepticism because it savored of perpetual motion. Brown had shown that it went on year after year and that it was exhibited even by inclusions in liquid droplets of geological age. It was commonly dismissed as due to conditions of observation such as one-sided warming or lowering of surface tension by the illumination, although it was found to be independent of the color and intensity of the light employed. Carbonelle and Thirion in 1874 explained it upon the basis of the kinetic theory as being due to the irregular impacts of the surrounding molecules of solvent.

Finally, the *psychological* difficulty was obviated by Einstein's [1] dictum, independently arrived at by Smoluchowski,[2] that from the standpoint of the kinetic theory each rigid particle is the exact equivalent of one molecule. Each has the same average kinetic energy $\frac{1}{2}mv^2$, and if m is large, v^2 is small, and vice versa. This accounts in the first place for the fact that the movements are greatest for the smallest particles, as shown in Figure 7·1, as adapted from Zsigmondy's observations.[3] Indeed, the smallest gold particles ever seen dart irregularly right across and out of the ultramicroscopic field.

A second important consequence of this point of view is that the actual path can never be seen in detail, since on the average it changes direction and magnitude one thousand million times per second.

Although the concept of such purely random motion is very simple, its mathematical discussion has commonly been very uncertain, being either unduly simplified or involving definitely erroneous assumptions,[4] such as that the motion is convergent and leads back to the place in which the particle was first noticed.

Here we need only Einstein's expression for the average total distance in a particular direction travelled by a particle in its random motion in time t:

$$A = \sqrt{t} \cdot \sqrt{\frac{RT}{N} \cdot \frac{1}{3\pi\eta r}}$$

where R is the gas constant, T the absolute temperature, N the Avogadro number (that is, the number of molecules in one gram mole), η the viscosity, and r the radius of an uncharged spherical molecule or particle.

The influence of temperature and viscosity was tested and confirmed by earlier workers. However, the classical investigation is that of Perrin.[5]

Historically, this investigation was of great importance because until this work on colloids and the almost simultaneous demonstration of the discontinuous nature of radioactive substances by means of the spintharoscope, many physical chemists questioned the real existence of atoms and molecules.
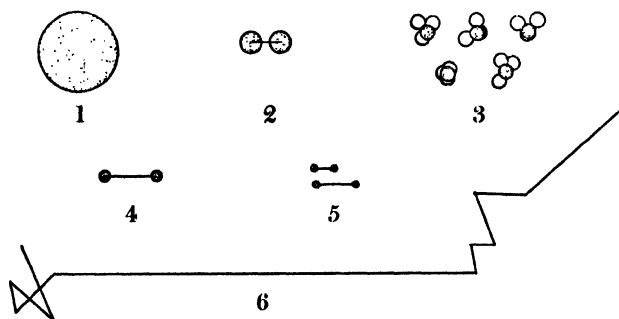
FIGURE 7·1 Brownian movement of particles of different sizes. After Zsigmondy, Text Ref. No. 3.

Perrin employed aqueous sols of gum gamboge and gum mastic prepared by throwing an alcoholic solution into water. These sols, like all others hitherto examined, contained particles of all conceivable sizes. By centrifuging the sols at a definite speed he threw out all the largest particles and discarded them. Then, with slightly increased speed he obtained the particles which had been just on the borderline

of sedimenting out, and these he diluted out again with water, thus obtaining unique monodisperse sols with very coarse particles of almost uniform size. Almost no one since has taken the trouble to prepare such monodisperse sols before making quantitative measurements with them.[6] His particles were chiefly 1 and 0.4μ in diameter. If Einstein's assumption were correct, these microscopic particles should exhibit a behavior resembling that of a gas with a molecular weight of several hundred thousand tons.

FIGURE 7·2 A drop of dilute sol of gum gamboge or mastic after evaporation. Taken from Perrin, J., *Die Atome*, p. 88 (see References).

For Einstein's equation it was necessary to know time, temperature, viscosity, and radius to determine N and to find whether N was a constant and of the magnitude expected. Hence, to obtain r it was necessary to know the amount of gum, the number of particles into which it was divided, and their density. The density of the precipitated particles was found to be the same as that of the original gum. They could be counted by sufficiently circumscribing the field of view so that the number visible at a given moment was small enough to recog-

nize, and the average of numerous observations was taken. This was checked after it was found that after precipitation the particles retained their individuality and spread out in sheets or rows upon a glass slide; see Figure 7·2, which is of interest in showing one mode of attachment of primary particles to form aggregates. By counting the particles, the radius could be confirmed by noting the number covering a measured area or the length of a row containing a definite number. It was also found to equal that calculated from the rate of fall using Stokes' law, equating the latter with the difference in weight of the particle and the water it displaced and multiplying by gravity g. If D is the density of the particle and ρ that of water:

$$6\pi\eta rv = \tfrac{4}{3}\pi r^3(D - \rho)g$$

Perrin plotted the position of a particle every 30 seconds, as in Figure 7·3. Thereupon he employed four methods of deriving N. First, he compared the amplitude of the movements in any given arbitrary direction with the theoretical values from the law of probabilities (see Table 6). The agreement remained even when the weight of the particle was changed 60,000-fold, and by additions of sugar, urea, and glycerin the viscosity changed 100-fold.

A second evaluation came from plotting each pair of successive observations with the first position as the origin, thus comparing the
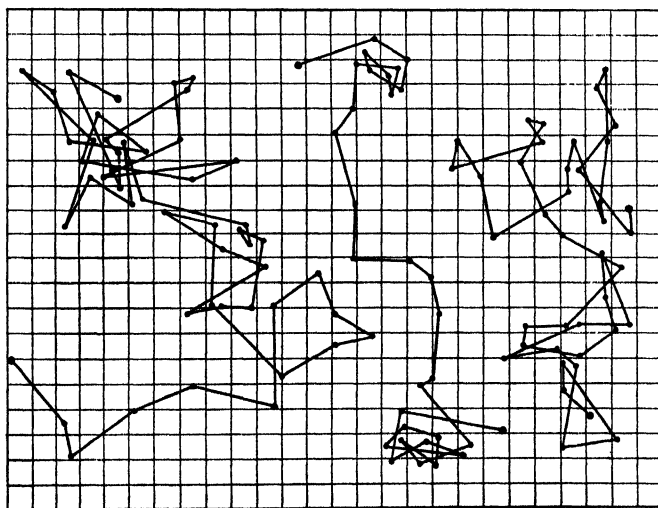
FIGURE 7·3 Observed position of a particle every 30 seconds. Taken from Perrin J., *Die Atome*, p. 107 (see References).

TABLE 6

Projected amplitude μ between	First Series		Second Series	
	No. Observed	No. Calculated	No. Observed	No. Calculated
0- 1.7	38	48	48	44
1.7- 3.4	44	43	38	40
3.4- 5.1	33	40	36	35
5.1- 6.8	33	30	29	28
6.8- 8.5	35	23	16	21
8.5-10.2	11	16	15	15
10.2-11.9	14	11	8	10
11.9-13.6	6	6	7	5
13.6-15.3	5	4	4	4
15.3-17.0	2	2	4	2

amplitude and the number of successive displacements, as shown in
Figure 7·4. The extent of agreement is shown in the following lines:

Calculated	32	83	107	105	75	50	27	14	7
Observed	34	78	106	103	75	49	30	17	9

A third evaluation of N came from observations of an irregular
large particle by equating the rotational with the translational energy.
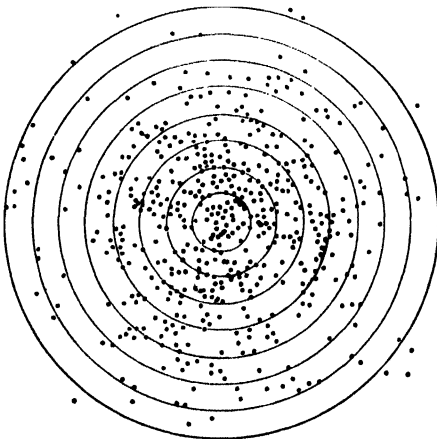
FIGURE 7·4 Direction and amplitude of movement of a par-
ticle of gum gamboge in 30 seconds, in successive observations.
Taken from Perrin, J., *Die Atome*, p. 110 (see References).

Finally, the value of N was determined by a very different method,
namely, the distribution with height in the upper few hundredths of a
millimeter in the surface of the sol. The distribution is logarithmic
like that in the earth's atmosphere where, for air of average tempera-
ture, the pressure is halved for each rise of 6 kilometers. In a sol of

uncharged particles, according to this law, n_1, the number of particles in a given small volume at a level of 1 cm will decrease to n_h at a level h cm higher, according to the formula of La Place:

$$\frac{n_1}{n_h} = 1 - \frac{N}{RT} m\left[1 - \frac{\delta}{D}\right]gh$$

where m is the mass of the particle, and the bouyancy of the liquid of density δ is taken into account. This formula is equivalent to:

$$2.3 \log \frac{n_1}{n_h} = \frac{N}{RT} m\left[1 - \frac{\delta}{D}\right]gh$$
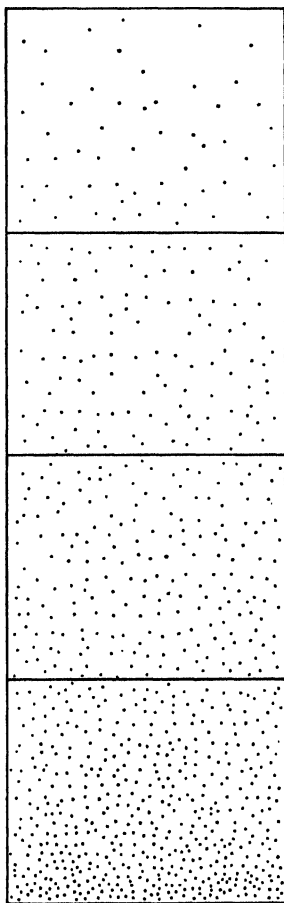
FIGURE 7·5 Weiser's diagrammatic reproduction of Perrin's *Die Atome*, p. 92. Taken from Weiser, H. B., *Colloid Chemistry*, p. 175, John Wiley and Sons (1939).

Thus in a Perrin sol with particles 0.1μ in radius, the formula predicts that the concentration should be doubled with every increase of 0.1 mm in depth. (In the last paragraphs of this chapter we discuss the fact that this does not occur except in the very uppermost layer.) Perrin's actual observations with a microscope focused to different depths are shown in Figure 7·5.

For N, Perrin thus found by the first two methods 69×10^{22} as compared with 68 and 65×10^{22} for the last two methods. The present accepted value is 60.227×10^{22}. Table 7,[7] which is not complete, lists the values for the Avogadro number which were established by 1915. This table illustrates one of the most important and dramatic achievements of science in linking together in quantitative manner most diverse and apparently unrelated phenomena through the use of this constant N. Thus an experimenter examining the Brownian movement of an emulsion under a microscope may predict the brightness of the blue sky or the viscosity of a gas or a rate of diffusion in a solution.

Of later work, it is interesting to note that Shaxby (1923) obtained 60.8×10^{22} from the Brownian movement of *Micrococcus (Staphylococcus) albus*. Sameshima (1925) and Houdijk (1928) with gum mastic and fine suspended wires respectively obtained 63.6×10^{22}.

TABLE 7

VALUES FOR THE AVOGADRO NUMBER UP TO 1915

BROWNIAN MOVEMENT	*Sedimentation equilibrium:*	
	Perrin (1909)	68.2×10^{22}
	Constantin (1914)	60
	Westgren (1915)	60.5
	Translation movement:	
	Perrin (1911)	68.5
	Svedberg (1911)	62
	Zangger and Böhl (1911)	61
	von Ettenreich (1912)	72
	Przibram (1912)	62.5
	Nordland (1914)	59.1
	Fletcher (1914)	60.3
	Diffusion:	
	Svedberg (1911)	58
	L. Brillouin (1912)	44
	Rotation movement:	
	Perrin (1909)	65
	Przibram (1913)	56
KINETIC GAS THEORY	*Internal friction and volume in condensed condition: (Loschmidt's Method):*	
	Mercuryvapor, Perrin (1909)	44
	Internal friction and van der Waals equation:	
	Air, van der Waals (1873)	60
	Argon, Perrin (1909)	62
	Helium, Ghose (1909)	62
	Theory of the liquid state:	
	Mercury, G. Jaeger (1911)	61.2
RADIOACTIVE MEASUREMENTS	*Charge on particles:*	
	Regener (1909)	60.4
	Rutherford and Geiger (1909)	62.4
	Helium production:	
	Dewar (1910)	57.3
	Boltwood and Rutherford (1911)	61.5
	Debierne (1914)	64
	Decomposition velocity:	
	St. Meyer (1913)	61
FOG METHODS	Lattey (1909)	62.1
	Millikan and Begeman (1910)	69
	Begeman (1910)	61.8–62.4
DROP METHODS	Millikan (1913)	60.62
	Roux (1913)	69
	Schidlof and Murzynowska (1913)	61.1
	Lee (1914)	60.7
PLANCK'S THEORY OF RADIATION	Planck (1913)	62.0
	Coblentz (1913)	60.8

Likewise, Brownian movement has been studied to determine N from the value of fluctuations in the concentration of particles in a very small volume, (a) as regards their magnitude and (b) as regards the rate at which they occur. From such a study Westgren obtained values which individually varied over rather wide limits, like those already discussed, but which gave a final average of 60.9×10^{22}.

After recording these considerable triumphs, it is necessary to mention two sources of misgiving as to the numerical accuracy of individual investigations with some of the colloidal sols studied. First, there is a conflict of evidence as to whether in certain gold sols the particles were completely separate and independent or partially aggregated. An error operating in the opposite direction was not noticed until 1925 or 1926; namely, that all the particles were charged, and, hence, each particle was accompanied in the solution by n other free charges. In the absence of any other electrolytes in the solution, this would give an apparent molecular weight only $1/(n + 1)$ times the true value. Only if the concentration of other electrolytic impurities was sufficient to act as a buffer would this error be reduced to negligible proportions, as is discussed in the chapter on diffusion.

For some years there was discussion and controversy over the fact that Perrin and others had observed sedimentation equilibrium or a logarithmic distribution of concentration only in a very thin layer on the upper surface of the colloidal sol, and it is obvious on inspection that colloidal sols, no matter how old, appear to be of uniformly distributed concentration throughout the body of the liquid. The difficulty, however, proved to be merely an experimental one of avoiding convection currents due to fluctuation of temperature. McDowell and Usher [8] showed that gold sol will establish its logarithmic concentration gradient over a depth of a centimeter if the temperature is not allowed to vary more rapidly than $0.001°$ per hour. A similar demonstration was made by Johnston and Howell.[9] Unfortunately, in both cases only the fact of logarithmic distribution was demonstrated, and the actual gradient of concentration was not checked against that which would have been predicted from a knowledge of the size of the gold particles.

Summarizing this chapter, we see that Brownian movement is just a visible manifestation of the kinetic theory of heat and that every independent rigid particle however small or large has the same average kinetic energy. This is equally true for a molecule as for the moving system of a galvanometer to which this fact sets a theoretical limit of sensitivity. The thermal movement of all molecules and par-

ticles is the basis of all such properties as diffusion, osmotic pressure, and the other colligative properties. For some very long, flexible molecules (cf. Chap. 20, High Polymers) the Brownian movement of the segments may take place semi-independently, and the colligative properties are correspondingly affected.

REFERENCES

Brownian Movement

DUCLAUX, J., *Mouvement brownien. Parte experimentale*, Hermann & Cie, Paris (1938)
PERRIN, JEAN, *Les Atomes*, F. Alcan, Paris (1913); *Die Atome*, Theodor Steinkopff, Dresden and Leipzig (1914); *Atoms*, D. Van Nostrand, New York (1916)

Text References

1. EINSTEIN, A., *Ann. Physik*, 17, 549 (1905); 19, 371 (1906)
2. SMOLUCHOWSKI, M., *Ann. Physik*, 21, 756 (1906)
3. ZSIGMONDY, R., *Zur Erkenntnis der Kolloide*, G. Fischer, Jena (1905)
4. For an illuminating criticism see Jaeger, *Sitzb. Akad. Wiss. Wien* (2a), 128, 1271 (1919) or for modern mathematical discussion see PALEY and WIENER, *Fourier Transforms in the Complex Domain*, Williamstown Colloq. Series, No. 19, Amer. Math. Soc., New York (1935) or DUCLAUX, *J. phys. radium*, (8)1, 81 (1940), and a review article by E. Kappler, *Naturwissen.*, 27, 649–666 (1939)
5. PERRIN, *op. cit.*
6. But see USHER, F. L., *Trans. Farad. Soc.*, 21, 406 (1925)
7. WESTGREN, A., *Z. anorg. Chem.*, 93, 240 (1915)
8. MCDOWELL, C. M. and USHER, F. L., *Proc. Roy. Soc.*, 138A, 133 (1932)
9. JOHNSTON, N. and HOWELL, L. G., *Phys. Rev.*, 35, 274 (1930)

ULTRAFILTRATION

Filtration through ordinary filter paper holds back all but the smallest particles visible in the microscope. Special filter papers and filter candles of porous earthenware or sintered glass extend the range slightly farther. Ultrafilters are those which are effective throughout the range of colloidal and even molecular sizes. The ultrafilter is now one of the most important means of studying colloid systems. An authoritative survey of this field has been published by Ferry.[1/]

Collodion had long been in use in bacteriology, but in 1907 Bechhold discovered that porous membranes of collodion could be precipitated by water from nonaqueous solutions and that the more concentrated the solution, the denser the resultant filter. He thus invented the ultrafilter of any desired porosity, except the very finest. His first papers,[2] still among the best on this subject, set forth many principles and applications. Elford [3] has standardized the procedure for making collodion membranes of fairly uniform pore size. The porosity of Elford's collodion membranes is surprisingly high, usually amounting to nearly nine tenths of the total volume, except for the finest membranes.

Field, in the author's laboratory, made collodion membranes which held back potassium chloride and sucrose. Kistler [4] found that collodion membranes can be made as dense as desired by filtering through them solutions of cellulose or of collodion. Many other less satisfactory materials have been used, such as gelatin hardened with formaldehyde,[5] animal membranes,[6] electrically deposited metals,[7] or a bed of graded carborundum particles.[8]

Commercial Cellophane (but not waterproofed), Sylphrap, and Viscacelle are widely available cellulose membranes which possess the

advantage of being chemically inert and insensitive. They have many pores fine enough to hold back a portion of some of the smallest molecules.[9] As a rule, Cellophane swollen in 64% zinc chloride solution allows all molecules to pass freely and holds back all colloids. Unfortunately, the use of cellulose membranes cannot yet be extended above the smallest colloid ranges. Collodion may be converted into more inert cellulose with only slight increase in density of the ultrafilter. All membranes may be used with any desired nonaqueous solvent, provided the membrane is first washed with a solvent in which water is miscible and then with a second solvent in which the first solvent is miscible.[10] Such a series is ethylene glycol with cellosolve, cellosolve acetate, gasoline, or kerosene.

FIGURE 8·1 Sectional drawing of metal ultrafilter, showing a membrane supported on a sieve. Taken from Barnard, J. E., and Elford, W. J., *Proc. Roy. Soc.*, **109B**, 376 (1931).
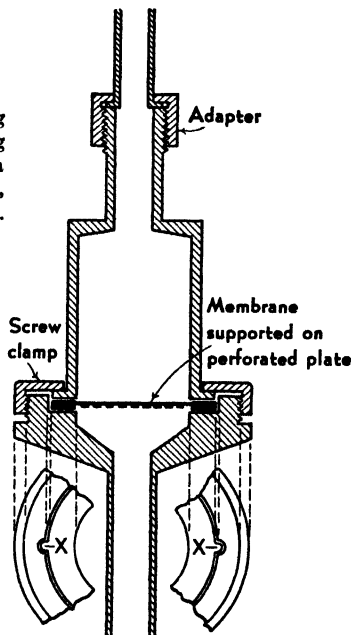
Most membranes require mechanical support, especially if, to save time or to overcome osmotic pressure, high pressures are applied. The membranes may be embedded in filter paper or in porous earthenware or unglazed porcelain, such as crucibles or evaporating dishes. They may then be used with a suction pump. They are frequently supported on fine sieves or filter paper or sintered glass or fine wire gauze (see Fig. 8·1), resting upon a stronger framework. For high pressure ultrafiltration the filter and its support are clamped between rubber or cellophane washers in a heavy metal casing such as was supplied

by Göttingen Commercial Filters,[11] who also supply ultrafilter membranes. It is advisable to replace the interior metal with a pure silver lining, or at least plate it with silver, platinum, chromium, or rhodium. Such filters may be attached directly to a gas cylinder supplying pressures over 100 atmospheres. It is often important to use a stirrer. That of the Göttingen filters is operated from the outside by a rotating electromagnet. An applied e.m.f. will hasten ultrafiltration in the same manner as in electrodialysis (Chap. 13).

Ultrafiltration must be distinguished from other related phenomena, such as diffusion, dialysis, osmosis, and electroosmosis. If two parts of one body of liquid are of different concentrations, the spontaneous process of equalization is called *free diffusion*. There is still the tendency for the solutes to diffuse into the less concentrated and the water into the more concentrated portions. As the pore size of the filter diminishes toward that of the coarsest particles, the diffusion of solute is impeded.

Dialysis is a differential diffusion where the membrane is permeable to molecules and ions but is too fine to allow colloid particles to pass. The former diffuse into the water while the water diffuses into the solution. In dialysis under pressure sufficient pressure is applied to offset the diffusion of the water, so that the concentration of the colloidal particles remains unchanged while the crystalloidal solutes diffuse through. In electrodialysis a current is applied to hasten the purification, the salt impurities around the electrodes being continually removed in flowing water.

In osmosis the membrane is sufficiently fine to allow only water to pass, whether through pores or by solution in the membrane. It is then called semipermeable. At osmotic equilibrium, the osmotic pressure balances the tendency for water to diffuse. As will be seen later, membranes exhibit an electrical potential different from that of solutions with which they are in contact, and this often causes a flow of liquid through the membrane in a direction opposite from that to be expected merely from concentration and osmotic pressure.

Ultrafiltration occurs where the applied pressure exceeds the osmotic pressure, or where, with a more permeable membrane, both water and diffusable solutes are forced through the membrane. In practice so much pressure is employed that diffusion effects tend to become negligible. During ultrafiltration the concentration of the residue must increase as the filtrate comes off. However, the process is sometimes complicated by sorption in the membrane, and this must always be taken into account. For example, Bechhold found that

arachnolysin (that of the spider), as is usual with intensely active poisons, is retained in the filter itself. This, like ultrafiltration, depletes the filtrate, but in ultrafiltration the liquid above the membrane steadily becomes more concentrated, whereas in sorption by the filter it too tends to be depleted. Hence it is more important to examine the residue above the filter than it is to examine the filtrate. For some purposes, examination of the latter may even be omitted.

The tendency of the membrane to react with the liquid will in general be most pronounced with the first layer of molecules or particles. This is indicated diagrammatically in Figure 8·2, representing three modes of filtration, in all of which the membrane sorbs a monolayer. "Normal filtration" occurs when the particles move unhindered through pores whose diameters after sorption are adequate. With surface active agents present, particles can go through smaller pores. "Abnormal filtration" is said to occur when the solid particles crowding toward the pore opening jam and tend to block the pore. "Blocking" is favored by high pressures, thick membranes, high concentrations, and absence of surface-active or stabilizing agents. It is connected with instability of the particles and seems to indicate their solid nature, because liquid droplets such as lecithin show an opposite behavior, high pressure sufficiently deforming them so as to force them through a hole of smaller diameter than the droplet. Thus, with pores slightly smaller than the droplet, high pressures allow droplets to pass through, while low pressures do not. In contrast, moderate pressures allow solid particles to pass through larger pores, but sufficiently high pressure may stop this even when the pore is ten times the diameter of the particles. Soap solutions exhibit blocking under suitable conditions and therefore are presumed to comprise solid colloidal particles, in conformity with direct x-ray evidence. Protective colloids often assist other particles to pass through, as would be expected from an inspection of Figure 8·2.
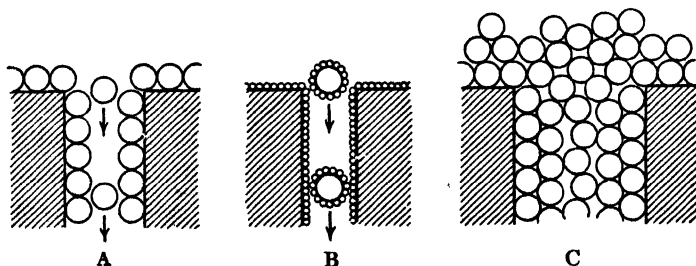
FIGURE 8·2 Schematic representation of normal and abnormal filtration. Taken from Ferry, Text Ref. No. 1.

The standard method of determining the size of the pores is to observe the pressure of air required to blow bubbles through the wet membrane.[12] It is evident that the resistance to be overcome before a bubble appears is dependent upon the surface tension of the liquid and that the first bubble will come from the largest hole. For water, 100 atmospheres is required to blow a bubble through a pore 284 Å in diameter, in accordance with Cantor's law:

$$d = 4\sigma/p$$

where σ is the surface tension and p the pressure. The pressure is inversely proportional to the pore diameter d, so that 10 atmospheres is enough for 2840 Å and 1 atmosphere for 28,400 Å. Knöll points out that absolute values are usually not obtained. Most of the filtration will occur through the group of larger pores. The size of pores can also be tested by the rate of flow of liquids. Another more convenient method is the test with a series of solutions whose particles are of known size, as determined by counting in the ultramicroscope or with previously measured membranes. Such a series of solutions of standardized mode of preparation is described by Krueger and Ritter.[13] Figure 8·2 shows the probable reason why a particle often requires a pore several times larger than itself for normal filtration.

The pores of a filter paper are about 1 micron (10,000 Å) in diameter; those of bacteria-tight earthenware filters are 1000 Å, and ultrafilters range from above this down to less than 2 Å. The ideal ultrafilter is a sieve. Liquids flow through membranes in accordance with Poiseuille's law just as if a membrane were equivalent to a bundle of parallel capillary tubes. Hence, the volume V flowing through in time t under a pressure p is:

$$V = \frac{n\pi r^4 pt}{8\eta l}$$

where n is the number of effective capillaries of radius r, and length l is equal to the thickness of the membrane, while η is the viscosity of the liquid.

Duclaux in 1924–25 found that the rate of flow of various liquids through a nitrocellulose membrane was proportional to the pressure and had the same temperature coefficient as the viscosity, where the viscosity of the different liquids varied by 200-fold. Collander [14] tested semipermeable membranes of copper ferrocyanide with solutions of twenty-five nonelectrolytes and eleven acids, finding the permeability inversely proportional to the molecular volume and the greatest pore diameter about 4 Å. He deduced, that the membrane consisted of a

ramifying aggregate of solvated colloid particles. It is evident from experience with the delicate copper ferrocyanide membranes that care must be taken not to alter the properties of the colloid membrane itself.

The ultrafilter has many uses. Bechhold used it not merely to sort out colloidal particles of various sizes but also to separate a coarser colloid, such as Prussian blue, from a finer one, such as hemoglobin. He also carried out fractional filtration of albuminoses. Sorption effects within colloidal solutions are readily studied by seeing whether or not the constituents filter independently of each other. For example, albumin in solution sorbs methylene blue and thus the two are not separable by filtration, as is indicated in Figure 8·2.

A sample of commercial Cellophane soaked in water only, instead of in 64% zinc chloride solution, was but partially permeable to ordinary molecules, as is shown in Table 8, taken from McBain and Stuewer,[15] where the proportion retained is parallel with the number of carbon atoms in the molecule. It may therefore be used to study molecular size. Cellophane that has not been swollen is practically gas-tight.

TABLE 8

PERCENTAGE OF DIFFERENT MOLECULES HELD BACK BY
ORDINARY CELLOPHANE

Solute	Initial Concentration Per Cent	Final Concentration Per Cent	Per Cent Retained
Glycerol (C_3)	2.96	3.02	2.2
Dextrose (C_6)	2.86	3.20	7.0
Sucrose (C_{12})	2.96	3.53	14.9
Raffinose (C_{18})	2.96	3.58	18.1

An important use of the ultrafilter is in the determination of the hydration of the colloid particles that are held back. A striking example is the filtration of a solution containing both potassium laurate (KL) and potassium chloride through a membrane which held back practically all the laurate but allowed potassium chloride to pass freely. As shown in Table 9,[16] the filtrate is 1.2 N KCl throughout, whereas the original solution was 1.0 N KCl, and the final residue 0.67 N KCl. In other words, the filtrate appears far more concentrated than the liquid being filtered. This can only mean that some of the water in the soap solution is definitely attached to the soap which is being held back and is not available for dissolving the potassium chloride. Hence, the real concentration of the potassium chloride in the solvent actually

TABLE 9

ULTRAFILTRATION OF A SOLUTION ORIGINALLY 1.0 N_w IN RESPECT TO POTASSIUM
LAURATE AND TO POTASSIUM CHLORIDE FILTERED THROUGH CELLOPHANE NO. 600,
AT 120 LB/SQ IN

| Residue | | Filtrate | | Hydration* |
N_w KL	N_w KCl	N_w KL	N_w KCl	moles H_2O/mole KL
1.000	1.000	0.01	1.194	10.3
1.088	0.9287	0.014	1.182	12.8
1.062	0.8544	0.01	1.203	12.7
1.287	0.9076	0.015	1.184	9.9
1.404	0.894	—	1.192	10.2
1.601	0.8341	—	1.206	10.8
1.767	0.8178	—	1.183	9.7
1.962	0.7955	0.05	1.200	9.8
2.28	0.718	—	1.19	9.6
2.584	0.670	—		

* Hydration in moles water/mole soap $= 55.5 \, (R_F - R)(R_F S - R S_F)$ where all concentrations are in moles. R is the mean concentration of the reference substance above the ultrafilter and R_F in the filtrate, S is the mean concentration of soap solution above the ultrafilter, and S_F that in the filtrate. The formula holds even when the soap is only imperfectly retained.

available is 1.2 N KCl throughout. Meanwhile, the soap (potassium laurate or KL) concentrates from 1 N KL to 2.6 N KL, its hydration corresponding to the formula KL·$10H_2O$.

The filtration of electrolytes is more complicated than that of nonelectrolytes, although even in the latter case there are distinct influences of pressure and rate of stirring. Very dilute electrolytes do not filter readily on account of the electrification of the membrane, which thus simulates the effect of small pores. This is obviated by the presence of any electrolyte in moderate or high concentration. It may be that the role of the salt added in dyeing is to restore effective pore size. Wool and silk and cellulose acetate are not permeated by dyes having large molecules,[17] which, therefore, tend to rub off unless the pores have previously been opened by suitable treatment or swelling.

Erschler[18] has convincingly pointed out that complete absence of stirring must soon suppress all minor differences between the composition of filtrate and original solution, because of the automatic building up of the compensatingly high concentration in the layer of liquid resting upon the upper surface of the filter. Extremely slow ultrafiltration approximates to the rate of dialysis, whereas very rapid ultrafiltration of solutions containing charged particles produces a filtrate of much lower concentration, as was shown with ferric hydroxide by McBain and McClatchie.[19] This should be borne in mind in connection

750mμ ⬤ **Bacillus Prodigiosus**

350mμ ⬤ **Spirillum Parvum**

250mμ ⬤ **Psittacosis Virus**

200mμ ⬤ ⎰ **Spirochaeta Pallida**
          ⎱ **Agalactia**

150mμ ⬤ ⎧ **Bov. Pleuropneumonia**
          ⎪ **Sewage Microorganism**
          ⎨ **Vaccinia**
          ⎩ **Canary Pox**

125mμ ⬤ ⎧ **Rabies**
          ⎪ **Infectious Ectromelia**
          ⎨ **Herpes**
          ⎩ **Pseudo Rabies**

100mμ ⬤ ⎧ **Influenza**
          ⎨ **Borna Disease**
          ⎩ **Newcastle Disease**

85mμ ⬤ **Vesicular Stomatitis**

75mμ • ⎰ **Rous Sarcoma**
       ⎱ **Fowl Plague**   ⬤ **75mμ**

        Phages (Staph "K", C16, D4, D12)   ⬤ **60mμ**

        Phages (D54, S41, Megatherium)   ⬤ **35mμ**

        Rift Valley Fever ────────⎫
        Equine Encephalitis─────── ⎬──⬤ **30mμ**
        Tobacco Mosaic ───────────⎭

        St. Louis Encephalitis──────⎫
        Phages (Typhoid TIII, C36, D13, D20, D48)⎭─⬤ **25mμ**

        Haemocyanin Helix
        Yellow Fever ───────────────⬤ **22mμ**
        Louping III ────────────────● **17mμ**
        Phage (S13) ──────────⎫
        Poliomyelitis ───────── ⎬─● **10mμ**
        Foot & Mouth Disease ──⎭

        Edestin ─────────────────────● **8mμ**
        Serum Globulin
        Serum Albumin ─────────⎫
        Oxyhaemoglobin ─────────⎬● **5mμ**
        Egg Albumin ─────────────• **4mμ**

FIGURE 8·3 From bacterium to molecule. Taken from Elford, W. J., *Trans. Farad. Soc.*, **33**, 1103 (1937).

with the use of ultrafiltration to determine the intermicellar liquid in colloidal sols. Further effects of the electrical properties of membranes on selective penetration by ions and causing flow of liquids through the membranes will be discussed in Chapter 14 on Diffusion.

Dean and Moore [20] have used nonprotein substances of high molecular weight to measure the pore size of the fertilization membrane in the eggs of the sand dollar. Concentrations of 0.2 millimole cause its collapse if the molecules are larger than 10,000 molecular weight; otherwise they merely penetrate it.

The most important field of application of the ultrafilter during the last decade has been its use in the hands of Elford and his collaborators to determine the size of various species of virus and bacteriophage in conjunction with confirmatory measurements made by the methods of microscopy with ultraviolet light, the ultramicroscope, ultracentrifugation, and diffusion. An example of his results is given in Table 10,[21] and some sizes so deduced are pictorially represented in Figure 8·3, where it may be remembered that one millimicron, mμ, equals 10 Ångström units.

TABLE 10

Substance	End-point of Filterability in Å	Probable Size by Filtration in Å	Size by Other Methods in Å
B. prodigiosus	7500	5000–7500	5000–10,000 microscopically
Bovine P. pneumonia			
Spheres	3500	1750–2500	2000–5000 } microscopically (Barnard)
Particles	2000	1000–1500	2000 }
Vaccinia virus	2500	1250–1750	2100–2300 centrifugal analysis (Bechhold and Schlessinger, 1931) 1700–1800 (max.) U. V. light photography (Barnard)
Infectious ectromelia virus	2000	1000–1500	1300–1400 U. V. light photography (Barnard)
Herpes virus	2000	1000–1500	
Gold sol (purple)	800	300–400	500–600 ultramicroscopical count
Phage C 36	650	200–300	Bechhold and Villa (1926), B. coli phage, 350, optical Hetler and Bronfenbrenner (1931), B. coli phage, 228–120 by diffusion
Gold sol (red)	400	150–200	200 ultramicroscopical count
Foot and mouth disease virus	250	80–120	
Oxyhemoglobin	100	30–50	5 ultracentrifugal analysis (Svedberg) 54 diffusion (Northrop and Anson)
Egg albumin	60	. . . .	43.4 ultracentrifugal analysis (Svedberg)

It will be seen from Table 10 that the results are well substantiated and that numerical data were thus given for the first time for this large and very important group of "filter passing organisms."

REFERENCES

Text References

1. FERRY, J. D., *Chem. Rev.*, **18**, 373–456 (1936), with 296 references
2. BECHHOLD, H., *Z. physik. Chem.*, **60**, 257 (1907); **64**, 328 (1908)
3. ELFORD, W. J., *J. Path. and Bact.*, **34**, 505 (1931); see also BAUER, J. H. and HUGHES, T. P., *J. Gen. Physiol.*, **18**, 143 (1934)
4. MCBAIN, J. W. and KISTLER, S. S., *J. Gen. Physiol.*, **12**, 187 (1928); *J. Phys. Chem.*, **33**, 1806 (1929)
5. BECHHOLD, *loc. cit.*
6. MCBAIN and KISTLER, *loc. cit.*
7. MANNING, J., *J. Chem. Soc.* (London), 1127 (1926); WARRICK, D. L. and MACK, E., *J. Am. Chem. Soc.*, **55**, 1324 (1933)
8. DUCLAUX, J. and AMAK, M., *Compt. rend.*, 205, 315 (1937)
9. MCBAIN, J. W. and STUEWER, R. F., *J. Phys. Chem.*, **40**, 1157 (1936)
10. MCBAIN and KISTLER, *loc. cit.*
11. Agents in the U.S.A., Pfaltz and Bauer, Inc., New York City
12. KNÖLL, H., *Kolloid-Z.*, **90**, 189 (1940)
13. KRUEGER, A. P. and RITTER, R. C., *J. Gen. Physiol.*, **13**, 409 (1930)
14. COLLANDER, R., *Kolloidchem. Beihefte*, 19, 72 (1924); 20, 273 (1925)
15. MCBAIN and STUEWER, *op. cit.*, p. 1164
16. MCBAIN, J. W., KAWAKAMI, Y. and LUCAS, H. P., *J. Am. Chem. Soc.*, **55**, 2762, 2763 (1933)
17. Cf. MORTON, T. H., *Trans. Farad. Soc.*, **31**, 262, 272 (1935)
18. ERSCHLER, B., *Kolloid-Z.*, **68**, 289 (1934)
19. MCBAIN, J. W. and MCCLATCHIE, W. L., *J. Am. Chem. Soc.*, **55**, 1315 (1933)
20. DEAN, R. B. and MOORE, A. R., *Biodynamica*. 6, 159 (1947)
21. ELFORD, W. J., *Proc. Roy. Soc.*, 112B, 405 (1933)

THE LYOTROPIC SERIES OF IONS

The behavior of many colloids is greatly influenced by the presence of electrolytes. Both cations and anions exert effects, and a series of each can be set up showing the order of their relative effectiveness. These series are known as the lyotropic series.

The lyotropic series of ions is of general significance in a large number of fields in which colloids do not occur at all, as well as in such behavior of colloids as salting out, swelling, gelation, and many other properties of colloidal systems. Loeb rightly emphasized that some of the earlier experiments with colloids had not taken into account the effect upon them of variation in pH. The importance of keeping pH constant for comparable experiments is even more evident when it is found that the effects of ions are often not only dependent upon pH, but that the whole series may be reversed in another range of pH.

It was the work of F. Hofmeister [1] in 1888 and 1890 on the concentration of salts required to salt out egg albumin that first called attention to this subject. His data are quoted in Chapter 19. Hence, the lyotropic series have often been referred to as the Hofmeister series. In descending order of effectiveness the Hofmeister series for the anions is:

citrate > tartrate > sulfate > acetate > chloride >
nitrate > chlorate > iodide > sulfocyanide,

and for the cations:

lithium > sodium > potassium > ammonium > magnesium.

These orders, though not exactly, are approximately the same for all phenomena and are of fairly general applicability.

Michaelis [2] pointed out that for a colloidal jelly "Kobalyi," which is entirely without electrical charge, "the effect of different electrolytes on the swelling is quite enormous, and does not depend on the valence number or sign of the charge, but follows the Hofmeister series." Lithium ions cause enormous shrinking, while sodium and potassium have a very small effect.

The fact that this difference in the behavior of different ions of the same valency is an inherent property of those ions rather than some peculiarity of colloids is evidenced by the many fields of physical chemistry in which the series are manifest, and yet no colloids are present. Such are: the solubility of gases in solutions of electrolytes; the solubility of nonelectrolytes, such as urea, quinone, and hydroquinone in water; the catalysis of chemical reactions by ions, as in the inversion of cane sugar or the hydrolysis of ethyl acetate; the effects of ions on the surface tension of water or on interfacial tensions; the absorption spectrum of water, and the reflection of infrared radiation; the surface potential of salt solutions; the displacement of the temperature of maximum density of water; the heats of hydration of the ions and their ionic radii and their standard free energies; the miscibility of organic liquids such as cyclohexane and methyl alcohol (note that this occurs in the absence of water); the compressibility of salt solutions; the viscosity of salt solutions; the solubility of amino acids in water; and the sorbability of the various ions by charcoal or methyl cellulose, or indeed by various colloidal particles. As an example, the following ions all raise the surface tension of water against air in the order given:

fluoride > sulfate > chloride > bromide > nitrate > iodide > thiocyanide.

Some of the manifestations of the lyotropic series with colloids are: swelling, for example that of starch granules, and the minimum temperature at which they form gelatinous pastes (Samec); the gelation of sols and liquefaction of gels; the effect of ions on the viscosity of sols, bentonite or barytes; the coagulation of sols; ion exchange with colloid particles; the variation in bulk of precipitates such as ferric oxide; and the permeability of semipermeable membranes and all plant cells to different ions. For example, the order of permeability of cations to a copper ferrocyanide membrane is:

ammonium > potassium and sodium > lithium > barium > strontium > calcium and magnesium.

Lest the reader get too definite an impression that the order of the ions is precisely the same for all phenomena that they influence, it should be pointed out that not only do exceptions occur, but that the order for any one case may change with concentration.

Much greater definiteness than would otherwise have been possible has been given to the consideration of the lyotropic series by the success of the Dutch school in attaching to the individual ions numerical values which serve to indicate the order in which their effects will be exhibited. This was initiated by Bruins [3] and carried through by Buchner.[4] It will then remain to establish the explanation for these properties. Pauli and Valkó [5] ascribed the series to the respective sizes of the anhydrous ions which would condition the electrostatic field strength in their immediate neighborhood and therefore determine their hydration.

Exceptions to the lyotropic series are not uncommon. Some have been noted by Docking and Heymann,[6] who find that the solubility of gelatin in solutions of all alkali metal chlorides is the same. Likewise, Eckfeldt and Lucasse [7] find displacements, and even inversions, in the lyotropic series of anions in salting out experiments using various solvents. They explain the salting out effect as due to three competing factors: the electrostatic effect, compound formation, and ionic polarization.

The Lyotropic Numbers

The quantitative characterization of lyotropic effects was carried through by Bruins in a very ingenious manner which enabled him to sort out at first sight the rather complicated phenomena of the salting out action of various ions. The manner in which this was done is illustrated by Figures 9·1 and 9·2, which show the salting out of agar and gelatin, respectively. For purposes of easy comparison Figure 9·1 has been replotted from several different sets of data of Buchner; Figure 9·2 is reproduced from the summary article of Voet.[8] These are portrayed in a manner which requires a little consideration. The diagrams are based upon the effect of progressively replacing sodium sulfate by sodium salts of other anions, doing this molecule for molecule. In the figures the ordinates are the molar concentration of the respective sodium salts other than sulfate, while the abscissae are the concentrations of the sodium sulfate. Thus the amount of sodium sulfate taken alone to cause salting out is in the neighborhood of 0.6 molar for both agar and gelatin. Sodium ferrocyanide 0.4 molar is as effective, and
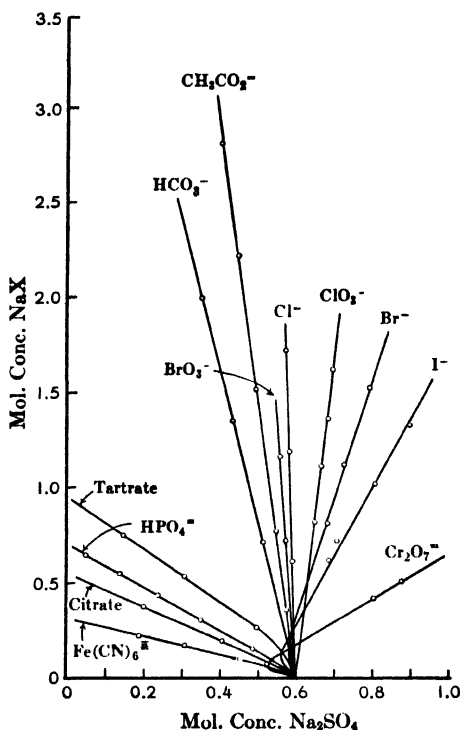
FIGURE 9·1 (left) Concentrations of different sodium salts required to flocculate agar-agar sols, plotted against the concentrations of sodium sulfate in the sols.
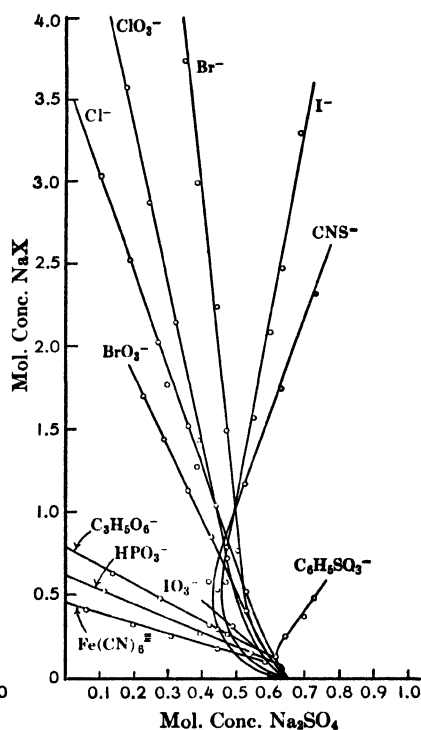
FIGURE 9·2 (right) Concentrations of different sodium salts required to flocculate gelatin sols, plotted against the concentrations of sodium sulfate in the sols.

intermediate mixtures of these two solutions are approximately as effective. Lines leading upward to the left belong to mixtures which help to salt out. A vertical line represents a salt having no influence and requiring a full amount of sodium sulfate throughout. Lines leading upward to the right are salts that salt in and have to be offset by an added dose of sodium sulfate. It will be noted that the lines corresponding to the initial admixtures with other salts are in general not straight and may often be of opposite sign to the slope exhibited by further additions. In other words, in most cases the initial addition of a salt helps to salt out even where further additions help to salt in. It will also be noted that whereas bromide salts agar in, it salts gelatin out. Nevertheless, Bruins[3] was able to establish the following regularity:

Bruins deals solely with the slope of the graphs after they have become straight lines and relates the slopes for gelatin and agar in the form of the linear equation:

$$\cot \alpha = A \cot \beta + B$$

where α and β are the angles which the lines make with the horizontal for gelatin and agar, respectively, and where A is a constant slope of this linear equation on a plot of $\cot \alpha$ against $\cot \beta$, and B is its intercept. Now it was possible arbitrarily to ascribe a number to each ion determined by the relation:

$$N = a \cot \phi + b$$

where a and b are constants. Since there are more constants than equations, an arbitrary value has to be inserted for one ion and the number $N = 2$ was selected for the sulfate ion and $N = 10$ for the chloride ion. Hence, for agar the equation was set up:

$$N = -4.00 \cot \phi + 10.00$$

and for gelatin:

$$N = -4.78 \cot \phi + 11.55$$

From this follows a corresponding number for each of the other ions. It is then found that the lyotropic numbers N for the respective ions agree well whether derived from the behavior of agar or gelatin, as is seen in the following table.

TABLE 11

THE VALUES OF THE LYOTROPIC NUMBERS N

Ion	BrO_2	NO_2	ClO_3	Br	I
Agar	9.72	10.1	10.74	11.50	12.50
Gelatin	9.38	10.2	10.58	11.14	12.48

For other ions the following numbers are derived:

F, 4.8; IO_3, 6.25; H_2PO_4, 8.2; NO_3, 11.6; ClO_4, 11.8; CNS, 13.25.

Armed with these numbers, one can test whether he can predict the salting out action of different salts in the case of some other colloid for which he has the action of two ions. If so, there should be a linear relation between the concentrations of those ions and the lyotropic numbers N. This is done in Figure 9·3 for the salting out of hemoglobin at pH 2.06, and it is seen that the linear relation does obtain. It is to be noted that here thiocyanate ion is the one most effective in salting

out so that the series is reversed in comparison with the data previously given in Figures 9·1 and 9·2 for agar and gelatin, but the relative order is maintained.

FIGURE 9·3 Concentrations of different sodium salts, necessary to flocculate hemoglobin sols, plotted against the lyotropic numbers.
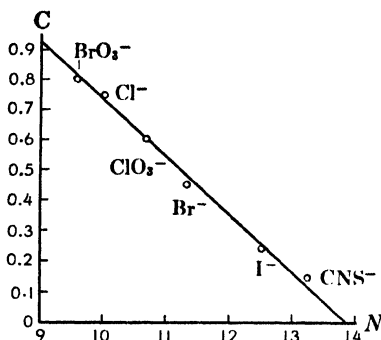
Figures 9·4 to 9·15 show the rather considerable success that has been achieved through the use of these lyotropic numbers in sorting out and even predicting the complicated behavior of the ions in various other cases. For example, in Figure 9·4 the results are given for the effect of salts of one fixed concentration, at 0°, upon the swelling of gelatin, for some shrink and some swell gelatin, as compared with the effect of water where the straight line crosses the abscissae at a value 9.8.

Figure 9·5 is a similar diagram at 18° instead of 0°, the value for water being again at 9.8.
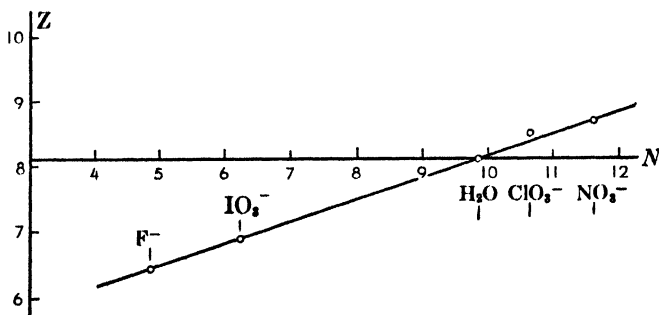
FIGURE 9·4 Relative weight increase of gelatin, swollen at 0° C., in solutions of different sodium salts of the same concentration, plotted against the lyotropic numbers.

FIGURE 9·5 Relative weight increase of gelatin, swollen at 18° C., in solutions of different sodium salts of the same concentration, plotted against the lyotropic numbers.
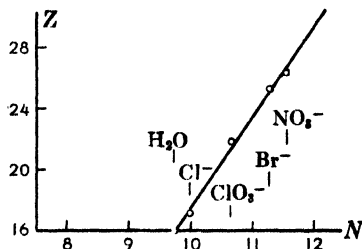
Figures 9·6, 7, and 8 show the lyotropic effects upon gelation, rate of reaction, and viscosity of electrolytes, all these being different regular linear relations. Linear relations are also found in the lyotropic series of cations in Figures 9·9 and 9·10, which refer to the concentrations of alkaline chlorides at different temperatures required to produce the same coagulation of gold and paraffin sols.
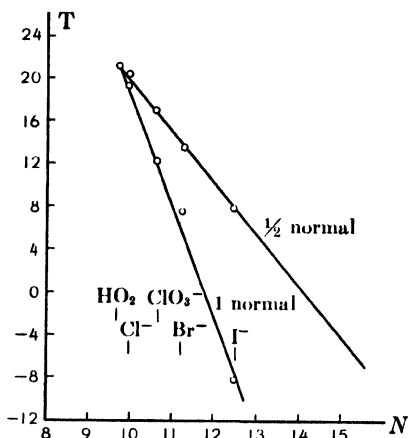
FIGURE 9·6 Gelation temperature of gelatin in solutions of different sodium salts, plotted against the lyotropic numbers.
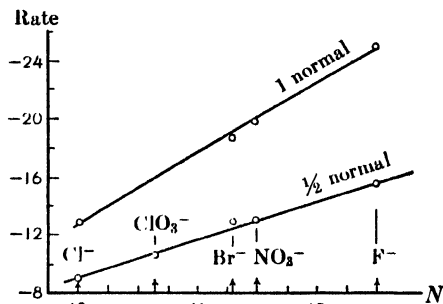
FIGURE 9·7 Rate of saponification of ethyl acetate as influenced by solutions of different sodium salts, plotted against the lyotropic numbers.
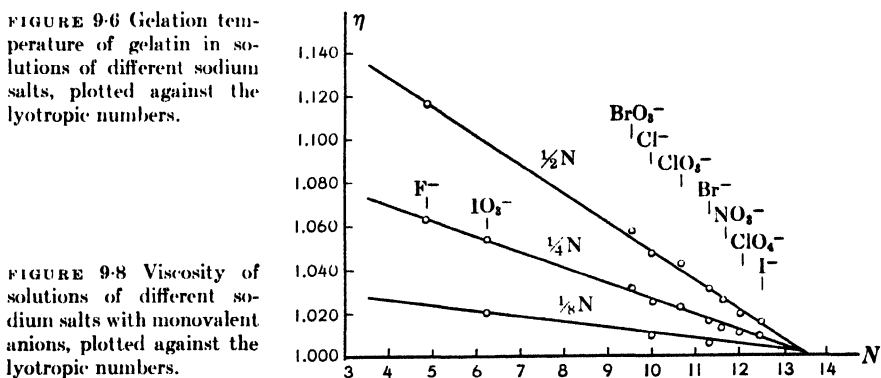
FIGURE 9·8 Viscosity of solutions of different sodium salts with monovalent anions, plotted against the lyotropic numbers.

Now for the cations it has been necessary again to ascribe lyotropic numbers, and the values of sodium and potassium were arbitrarily taken as 100 and 75, respectively, from which those for lithium, rubidium, and cesium then become 115, 69.5, and 60, respectively. So far, all the relations have been linear and this also holds for ion exchange. In Figure 9·11 the effects of the lyotropic series pass through a maximum; in Figure 9·12 through a maximum and a minimum. Without the guidance of the lyotropic numbers the data would have appeared completely irregular, particularly since the curve for dialyzed dextrin shows a minimum instead of a maximum.
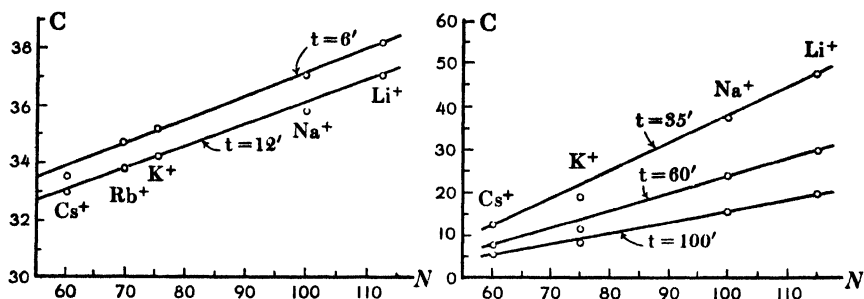
FIGURE 9·9 Concentrations of different alkaline chlorides, causing the same decrease in the number of particles of a gold sol in the same time interval, plotted against the lyotropic numbers.

FIGURE 9·10 Concentrations of different alkaline chlorides, causing the same decrease in the number of particles of a paraffin sol in the same time interval, plotted against the lyotropic numbers.

Now it is significant that the heats of hydration of the cations and of the anions, as plotted in Figures 9·13 and 14, again are in linear proportion to the lyotropic numbers. So are the ionization potentials, as shown in Figure 9·15. From these data we can calculate the heats of hydration of complicated ions which have not been directly measured or heats of hydration could be used as plotted instead of lyotropic numbers. Cooper [9] has introduced into their proper place in the series a large number of other ions from their Gibbsian standard free energies of formation from the elements.

It is clear that the earlier conceptions of Pauli and Valkó, that the ionic radius conditions the electrical field strength and therefore affects the hydration or solvation of the ions, are borne out by the data on the lyotropic numbers.

In addition to the significance of the heats of hydration, Bruins has discussed the effect of the charge on the colloidal particle and that on the ion and their effects in the orientation of the respective molecules of hydrate water. These mutual interactions lead to flocculation and stabilization, respectively, and show clearly why the lyotropic
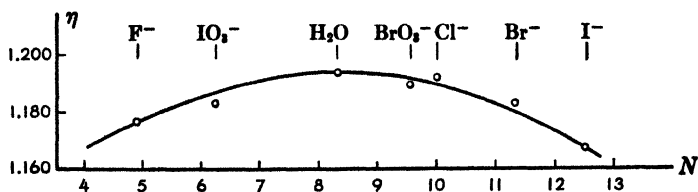
FIGURE 9·11 Viscosity of dextrin in solutions of different sodium salts, plotted against the lyotropic numbers.

series can be reversed, when, for example, the charge on a protein varies from positive to negative.

Voet [10] has uncovered linear relations between the ionic radii and the reciprocal of the heats of hydration for uni-, di-, and trivalent ions. High heats of hydration correspond with low numbers. These have high salting out power for negative sols and have the greatest effect in increasing viscosity or in causing gelation or in suppressing solubility. Both positive and negative ions with low lyotropic numbers produce the highest surface tensions and shift the maximum density of water to higher temperatures. Ions orient the water dipoles around them, and this is important when the sorbability of the ions plays a role. Thus, for a positively charged sol of egg albumin, ions with low lyotropic numbers are less effective in salting out because negative
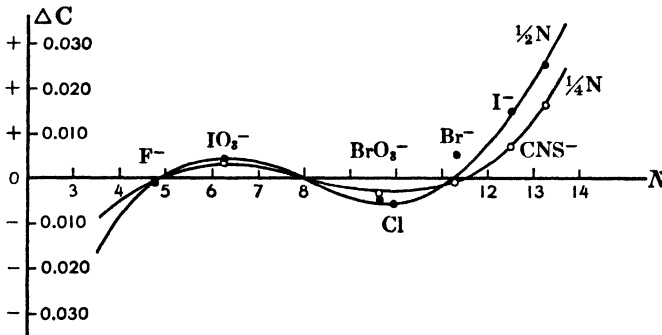
FIGURE 9·12 Adsorption of sodium salts by potato starch, plotted against the lyotropic numbers.
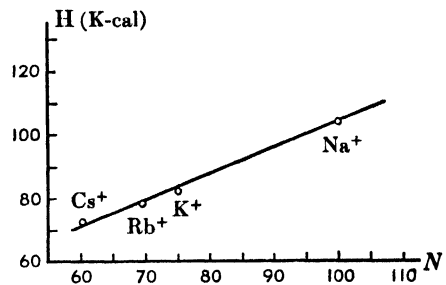
FIGURE 9·13 Heat of hydration of alkaline ions, plotted against the lyotropic numbers.

ions orient the dipoles with the negative end out, favoring their sorption by the positively charged colloid, and this sorption stabilizes the sol.

Again, in the coagulation of negative sulfur sols by positive ions a high heat of hydration makes for highly sorbed ions and poor flocculation. Finally, in the peptization of globulin proteins low lyotropic numbers or high heat of hydration gives a favorable orientation, good sorbability, and good peptizing action.

In seeking a more satisfactory explanation of the lyotropic series of ions, Buswell, Gore, and Rodebush [11] studied the effects of additions of ions of various radii and various energies of interaction on the infrared absorption bands of water. They showed that the effect of the halogen anions upon the molecular absorption coefficient of water is in a lyotropic sequence parallel with the respective halogen ionic radii. They postulated two opposing effects, one the depolymerization of the liquid structure and the other the electrostatic orientation of water molecules around the dissolved ion.
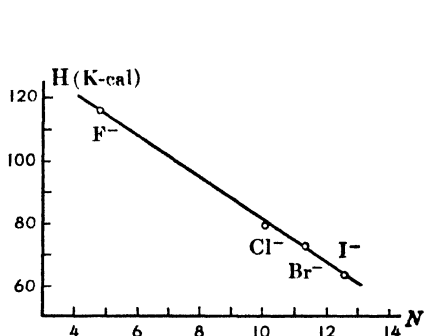
FIGURE 9·14 Heat of hydration of halide ions, plotted against the lyotropic numbers.
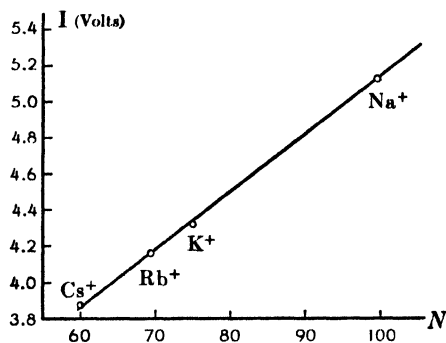
FIGURE 9·15 Potential of ionization of alkaline metals, plotted against the lyotropic numbers.

REFERENCES

Text References

1. HOFMEISTER, F., *Arch. exp. Path.*, 24, 247 (1888); 27, 395 (1890)
2. MICHAELIS, L., *Colloid Symp. Monog.*, 11, 14 (1925)
3. BRUINS, E. M., *Proc. Acad. Amsterdam*, 35, 107 (1932)
4. BUCHNER, E. H., *Rec. trav. chim.*, 53, 288 (1934); *Kolloid-Z.*, 75, 1 (1936); see also VOET, A., *Chem. Rev.*, 20, 169 (1937)
5. PAULI, WO. and VALKÓ, E., *Kolloidchemie der Eiweisskörper*, 2nd ed., Theodor Steinkopff, Leipzig (1933), p. 120
6. DOCKING, A. R. and HEYMANN, E., *J. Phys. Chem.*, 43, 513 (1939)
7. ECKFELDT, E. L. and LUCASSE, W. W., *J. Phys. Chem.*, 47, 183 (1943)
8. VOET, *loc. cit.*
9. COOPER, L. H. N., *Nature*, 139, 284 (1937)
10. VOET, A., *Trans. Farad. Soc.*, 32, 1301 (1936)
11. BUSWELL, A. M., GORE, R. C., and RODEBUSH, W. H., *J. Phys. Chem.*, 43, 543 (1939)

VISCOSITY, THIXOTROPY,

AND PLASTICITY AND DILATANCY

VISCOSITY

Viscosity is one of the most obviously important properties of colloid systems. Under various guises it is used in the control of very many industrial processes. For one thing, it varies quite enormously. Mixtures of turpentine and rosin, for instance, vary from a very mobile fluid to a plastic or elastic and brittle solid, the viscosity changing by more than 10^{20}. Besides, viscosity is something that can always be measured, though perhaps crudely, as by cutting with a knife or stirring with a paddle. When it is measured under constant conditions, a definite numerical value can always be obtained even though this may vary greatly with the method employed and its theoretical significance may be a matter of individual interpretation.

From a scientific standpoint, the behavior of colloid systems when made to flow, that is, their viscosity, apparent viscosity, plasticity, and elasticity, yields considerable insight into the structure of the system, the sizes of the particles, and the relation of the particles to each other. Although many aqueous sols, like gold, flow almost as readily as water, Graham was so impressed with the significance of the much higher viscosities frequently encountered with colloids that he called viscosity the "colloidoscope."

The subject of viscosity has an enormous although scattered literature. A *Journal of Rheology*, now included in the *Journal of Applied Physics*, has been devoted to this subject since 1930. Older summaries are the books by Bingham, 1922, and Hatschek, 1928, and that of Barr, 1931, while Houwink, 1937, reviews the behavior chiefly of the

more solid forms of colloids, such as glass, resins, asphalt, rubber, cellu-
lose, proteins, clay, and sulfur. The committee for the study of viscosity,
of the Academy of Science at Amsterdam, issued its first very valuable
and comprehensive report in 1935 and its second in 1938.

Definition

Following Sir Isaac Newton, we define the coefficient of viscosity
as the numerical value of the tangential force on unit area of either
of two parallel planes at unit distance apart when the space between
these planes is filled with the fluid in question and one of the planes
moves with unit velocity in its own plane, relative to the other. For
such a Newtonian liquid the flow is lamellar, each intermediate plane
of the liquid moving past its neighbors with the same velocity of shear.
Such flow is called truly viscous or Newtonian, and the force F neces-
sary to produce it is proportional directly to the velocity v and inversely
to the distance apart x of the planes:

$$F = \eta \cdot \frac{v}{x}$$

so that η is a proportionality constant and is then the force on 1 sq cm
of the plane moving with a velocity of 1 cm per second 1 cm away from
and parallel to the fixed plane. The quantity η is then the coefficient
of viscosity of the liquid in question. The unit of viscosity is the poise,
named after Poiseuille. It is useful to remember that the viscosity of
pure water at 20° C is very nearly 0.01 poise or 1 centipoise. Poiseuille
carried out a remarkable and extensive investigation in 1842 on the
rate of flow of water through forty different capillaries of various
lengths and diameters under various conditions with the object of
throwing light upon the circulation of the blood. He discovered experi-
mentally Poiseuille's law, which is also readily deduced from the defi-
nition of viscosity given above:

$$V = \frac{\pi g p r^4 t}{8 \eta l}$$

where V is the volume flowing in time t, g the acceleration due to
gravity, p the pressure in grams per square centimeter, r the radius
of the capillary tube, and l its length. All the terms are to the first
power, with the notable exception of the radius, r. In such purely
viscous flow, the movement of the liquid is devoid of turbulence and
is lamellar like the pulling open of a telescope, any plane in the liquid
normal to the tube becoming a paraboloid of revolution and the liquid
that actually wets the walls of the tube not moving at all.

It should be mentioned at once that only pure liquids and crystalloidal solutions and a limited number of colloidal systems exhibit a true coefficient of viscosity independent of the conditions of flow and the rate of shear. Nevertheless, resistance to flow under any given conditions is expressed for convenience in terms of viscosity.

Methods of Measurement

There are various methods of measurement and most of them do not lead to a result which, even for a liquid exhibiting true viscosity, could be expressed in absolute units or poises. However, it is almost as easy to use a method giving results that can be compared with those of other workers.

The commonest method used with freely flowing transparent liquids is the capillary viscometer of Wilhelm Ostwald. It consists essentially of a pipette with a capillary leg which is bent around and widened out to form a lower reservoir and then continued up to the same level as the top of the pipette so that the whole may be almost completely immersed in a thermostat. Figure 10·1 shows the general form of the glass tubing. By choosing the dimensions suitably in accordance with standard tables [1] all corrections for departures from Poiseuille's law are eliminated. The viscometer is standardized by first measuring the time taken by the surface of water or other known liquid, such as aqueous sucrose solutions standardized by the Bureau of Standards, to flow from the mark above the upper bulb to the one just below it. The ratio of the viscosities of the two liquids is directly proportional to the time required for each.

Many variants of the capillary viscometer are in use, some being rearranged for opaque liquids by placing the reservoir uppermost and reversing the flow. The most valuable are
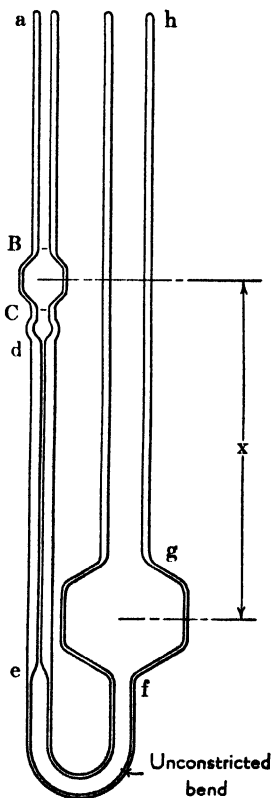
FIGURE 10·1 U-tube viscometer adapted for pipette filling. B and C are etched markings and should extend round the tube. The distance x is 10 cm. The markings a, d, e, f, g, and h are letters employed to define standard dimensions. From *Determination of Viscosity in Absolute Units*, p. 33 (see References).

those where the pressure head may be arranged to have any desired constant value so as to ascertain whether and how far the liquid departs from true viscosity.

More viscous liquids are measured with a falling ball method, such as is illustrated in Figure 10·2, where the time of fall of a steel ball between marks A and B on a wide test tube containing the liquid is compared with that observed with a standard liquid in the same tube.
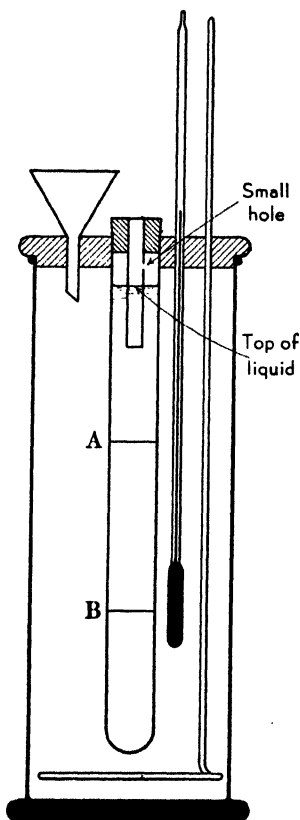
Again, the viscosities are proportional to the times. The commercial Höppler viscometer is a precision form with inclined tube which measures even gases and distinguishes between distilled and tap water, covering the range 10^{-2} to 10^5 poises.

Wherever a large, heavy ball does not fall or does not fall with sufficient rapidity, the whole apparatus may be mounted on a centrifuge [2] and again much valuable information is obtained by varying the centrifugal force.

The best design of instrument for careful, scientific investigation is the Couette type of viscometer, where one cylinder rotates within another fixed cylinder, the liquid filling the annular space between them. Such an instrument is shown in Figure 10·3. Here the attempt is made to confine all the friction to the known annular volume of liquid by putting in nonrotating end pieces above and below. Although much time and thought have been devoted to this purpose, no perfect design for accomplishing it has yet been achieved. Of the common commercial forms of such rotating viscometers, the MacMichael comes nearest to yielding true viscosities.

For semisolid colloids a form of apparatus is often used where under a known hydrostatic pressure the material is extruded through an orifice or a capillary tube. Deviations from the parabolic distribution of velocities corresponding to purely viscous flow may be detected by the shape of the extruded specimen or by inspection of a portion colored before distortion.[3] For semisolid or solid systems a penetrometer method is used, in which a needle or tube is pressed into the

FIGURE 10·2 Standard falling ball viscometer. From *Determination of Viscosity in Absolute Units*, p. 32 (see References).

material. For still more rigid specimens the resistance to torsion or bending of the specimen itself may be studied.

Once more it should be pointed out that the results of the different methods do not always agree within several fold and that they are often dependent upon the speed of rotation; the previous treatment of the system; how much working, flow, or vibration it has had; and how long it has been allowed to rest.

FIGURE 10·3 Couette type of viscometer. A hollow cylinder A, which is open and beveled at the ends, is suspended by a wire B from a heavy bracket C, provided with leveling screws (not shown). This cylinder A is coaxial with an outer cylinder D, which is provided with a water jacket E covered with insulating material. Two short guard cylinders, F and F' — closed at one end — are placed opposite the open ends of the cylinder, leaving a clearance of about 2 mm between the beveled surfaces. The outer cylinder D rests on a table G, which can be rotated by suitable gearing. The suspended cylinder A is provided with a mirror H, and its deflection read by scale and telescope. Taken from Hatschek, Text Ref. No. 8.
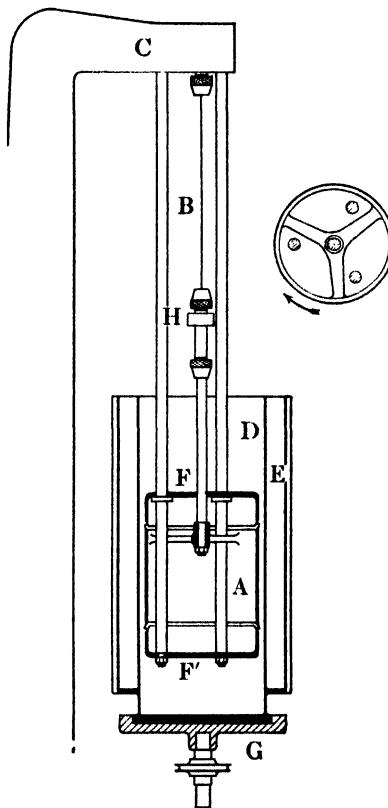
In general, the unstable sols of colloidal metals or sulfides have a viscosity not very much greater than that of the solvent. On the other hand, the stable sols, which can be prepared in much higher concentration, show much greater viscosity so that moderately dilute nitrocellulose solutions may be as viscous as cold molasses. Their apparent viscosity rises very rapidly with increasing concentration. Sometimes their temperature coefficient is so high that within a narrow range they seem to pass from nearly solid to fluid, almost like a melting, over a short range of temperature.

The viscosity of some of these aqueous sols is greatly affected by electrolytes. For instance, sodium sulfate increases the viscosity of gelatin solutions, while chlorides and nitrates diminish it. The pH, or acidity, is often particularly important. In some cases the effects of a given addition are not at all simple. For example, addition of sodium hydroxide or sodium chloride to a soap solution at first slightly lowers the viscosity, while further additions increase it enormously; but after attaining a maximum, further additions rapidly decrease it again until at a final concentration the soap separates into two layers.[4]

Theory of Relative Viscosity [5]

The only well-grounded formulation of viscosity is that of Einstein,[6] who applied the ordinary formulas of hydrodynamics to Brownian motion, viscosity, and diffusion of colloidal particles. However, he assumed that the particles were spherical and large as compared with a molecule of solvent, that they were uncharged, that there was no slip between them and the solvent with which they were actually in contact (that is, the solvent adhering to their surface), and, finally, that turbulence was avoided. He did not hesitate to apply his calculations to molecules of sucrose in water, relating the viscosity and diffusion to the Avogadro number N, which he thus deduced as being equal to 65.6×10^{22}. By this procedure he obtained the well-known formula for relative viscosity, η_r or η/η_o, where the viscosity η is referred to that of the solvent η_o as unity. If v is the fraction of the total volume occupied by the suspended molecules or particles, Einstein's formula is:

$$\eta_r = 1 + 2.5v$$

FIGURE 10·4 Cross section through suspension of spheres in a Newtonian fluid, showing that the fluid must undergo more shear even though the spheres minimize this by rotating. Spheres one third of the diameter shown would correspond more nearly to the example given in the text.

It should be noted that this theoretical viscosity is independent of the degree of subdivision, or size of the particles, as long as the latter are separate, uncharged spheres. Figure 10·4 shows how the presence of suspended particles makes a slightly increased demand upon the movement of the solvent between the particles. Thus, a solution of 0.1% by

volume should have a relative viscosity equal to 1.0025 times that of the solvent. Actually, many colloidal solutions of this low concentration exhibit an increase above the viscosity of the solvent many hundred times greater than this, as is, for example, the case for nitrocellulose.

The nearest to numerical agreement with the Einstein formula are dilute solutions of gum gamboge or of Sven Oden's sulfur, where, however, the constant has to be taken empirically as 2.9 or 3 instead of the theoretical 2.5. With more concentrated sulfur sols the discrepancy is greater, requiring a value of the constant 5.4 and 10 for 25% and 50%, respectively. For 1% egg albumin at its isoelectric point, Mrs. McBain found that the value 10 is required. Sachs in 1932 found that gold sols required a value of about 35 instead of 2.5. Carborundum sols required 19.

It seems that observed viscosities are at least somewhat higher than that predicted by Einstein.[7] In a few simple cases Hatschek pointed out that this could be accounted for by assuming an adherent rigid coating of solvent about 9 Ångström units thick (Cf. Chap. 4.). He supported this by showing that even in such a chemically indifferent system as starch suspended in toluene, the viscosity coefficient decreases with enhanced rate of shear as if some of the adhering envelope were rubbed off.

HYDRATION OR SOLVATION

In many cases, such as with soap, rubber, proteins, and derivatives of cellulose, the viscosity observed is enormously greater than that expected from the Einstein equation. Four explanations have from time to time been put forward; namely, solvation, electroviscous effect, effect of threadlike particles, and structural viscosity arising from the loose clumping or adhesion of particles to form ramifying aggregates immobilizing the solvent.

Solvation, whether due to chemical or physical forces, would have to be additional to, or different from, Einstein's adherent water and would not include enmeshed or enclosed solvent. Solvation would increase the volume of the particle. However, it would have to amount to a quite improbable extreme if it were to account for viscosities commonly observed. For 3% (by weight, 1.8% by volume) nitrocotton, where the viscosity may be 1000 times that of the solvent (instead of Einstein's only 1.0362-fold), it would be necessary to assume that the volume of the particles had increased to 99.8% of the total volume of

the solution, leaving 0.2% for the whole of the free solvent. However, direct experiment shows that particles sorb not more than their own weight of solvent vapor, which seems to set an upper limit to the possibilities of this once popular explanation, so that the effect of solvation, although real, accounts for only 1 or 2% of the excess viscosity.

Since this hypothesis would have required 3% nitrocotton to be an excessively concentrated emulsion consisting of the two immiscible phases, pure solvent and 97% solvent (!), we may digress for a moment to consider the formulation of viscosity of emulsions. When the emulsion is so concentrated that the particles crowd each other, a cross section of the emulsion at rest would, according to Hatschek,[8] be something like the hexagons shown in Figure 10·5A, each dodecahedron being surrounded by a continuous medium, shown in black. It is evident that such an emulsion would not flow at all but would show only an elastic yielding to stress until the droplets were still further deformed into practically continuous layers, as shown in Figure 10·5B. Then the flow would be confined to the horizontal layers of continuous black. For the resulting flow Hatschek deduced the formula:

$$\eta/\eta_o = 1/(1 - \sqrt[3]{v})$$

For blood corpuscles this formula holds within 1% up to an emulsion 90% by volume, where the viscosity has risen to 30.5, a result which is, however, far lower than that for the 3% nitrocotton already cited. It would hold only, of course, for deformable liquid droplets, not for solid colloidal particles.

FIGURE 10·5 Cross section through a close-packed concentrated emulsion with (A) droplets distorted to dodecahedra when at rest, and (B) elongated in parallel layers when the whole is forced to flow.

Many empirical formulas have been proposed to describe approximately the behavior of viscous colloid solutions. Many of them are modifications of the empirical Arrhenius equation:

$$\log \eta/\eta_o = k \cdot c$$

where k is a constant and c the concentration, and the graph should be a straight line instead of the distinct curve usually obtained.

ELECTROVISCOUS EFFECT

Von Smoluchowski [9] long ago pointed out that since colloid particles in general are electrically charged, their movement must meet with some added resistance, which would increase the viscosity predicted from the Einstein equation for uncharged spheres to the following extent:

$$\eta/\eta_o = 1 + 2.5\ v\left[1 + \frac{R}{r^2\eta}\left(\frac{D\zeta}{2\pi}\right)^2\right]$$

where R is the specific resistance in electrostatic units, D the dielectric constant, ζ the electrokinetic potential (discussed in Chap. 13), r the radius of the particle, and η the viscosity of the sol. For example, if for an aqueous sol with conductivity about that of ordinary conductivity water, $R = 1.1 \times 10^{-6}$ mho and $D\zeta = 81 \times 0.038$ volts = 3 volts = 0.01 statvolt, and $\eta = 0.01$ poise, and $r = 1500$ Å, the effect would be just large enough to change the constant 2.5 to 5 and therefore double the small difference in calculated viscosity between sol and solid. The effect vanishes completely at the isoelectric point and likewise disappears if there is even a very small amount of electrolyte present, as in distilled water. Hence, the effect again is real but very small. Although some writers still stress the electroviscous effect, it was found that in a typical case the calculated effect was only one hundredth part of the allowance made for the effect of hydration which has already been discussed as a real but minor factor.

EFFECT OF EXTREME ELONGATION OF PARTICLES

Von Smoluchowski likewise pointed out that Einstein's calculation for spheres would require a larger constant for particles departing greatly from that shape. In recent years Staudinger has explored very systematically the possibility of ascribing the high observed viscosities of many solutions of natural synthetic products to this source. The synthetic products are the artificial polymers to be discussed in Chapters 20 to 25. It is assumed that many, like cellulose derivatives, consist of single extremely elongated rodlike molecules. He then formulates the viscosity as follows:

$$\eta_{sp}/c = K_m \cdot M$$

where $\eta_{sp} = \eta/\eta_o - 1 = \eta_r - 1$, M is the molecular weight, K_m is a constant of the order of 1×10^{-3} at 20° depending somewhat on the nature of the homologous series, and c is the concentration per liter

in terms of the repeating links in the long chain molecule. For example, for paraffin one link would be CH_2, weighing 14 gm. One of these per liter would be a 1.4% solution and c would be 1.4. The application of the formula is restricted to solutions so dilute that η_{sp} is not more than 0.1–0.3. The solution should be so dilute that η_{sp} is still proportional to the concentration. If Arrhenius' formula still holds, this retains an error in η_{sp} of 5 and 13% for value of η_{sp}, equal to 1.1 and 1.3, respectively.[10] The ratio η_{sp}/c extrapolated to infinite dilution is commonly used and is called "intrinsic viscosity" and is given the symbol $[\eta]$.

The value of K_m is obtained empirically by a process of extrapolation for molecules so small that the molecular weight may still be determined by the usual physical-chemical methods; the value of K_m is deduced by comparison with the observed viscosity, and it is then assumed that the formula will apply to the much larger molecular weights in question. The procedure has the merit of yielding a definite molecular value for M, whether or not it is correct. The extrapolation necessarily includes any other source of progressive increase of viscosity, such as association or agglomeration still to be discussed. Likewise the measurable lower members are, as Lawrence has pointed out, hardly rodlike, when for a straight chain polyoxydecoic acid of molecular weight 780 the length is 20.6 Å, while the width is 13 Å and the thickness 5 Å.

Likewise, the author and his wife have pointed out [11] that the pure hydrodynamic formulas involved [12] yield an effect of a lower order of magnitude, indicating that Staudinger's effect is real but still minor, like solvation and the electroviscous effect. For a time Staudinger's hypothesis came into great prominence, but in recent years many authors have pointed out severe restrictions. Dobry [13] found that different fractions of polysterol, although of constant molecular weight, gave different viscosities. McBain and Scott [14] found that the constant pure geometrical effect demanded by Staudinger changes many fold between 0° C and 25° C with the same solution, that different solvents yield different results, and that sometimes the viscosity may be increased or decreased by simple treatments that cannot alter the molecular weight. It is now restricted to infinitely dilute solutions, and the constant differs for different solvents using the same substance, and the value of the molecular weight has been shown by other methods to come out too high. Currently, the Staudinger formula is modified to $[\eta] = K \cdot M^a$, where a has a value in the neighborhood of 0.67 but varies distinctly with the solvent. For example, for polystyrene in

cyclohexane it is 0.71, in carbon tetrachloride 0.61, in benzene at 60° C it is 0.65 but in benzene at 25° C only 0.53.

It is evident and is now generally accepted that long, more or less stringlike molecules like the cellulose derivatives just discussed, and some other high polymers, do not behave as long, rigid particles. Instead they curl up more or less into a tangle, depending upon their relation to the solvent. This, of course, partially immobilizes some of the solvent within the random coil or tangle.[15] However, as will be shown in later chapters, many high polymers show totally different behaviors. Thus proteins are solid bodies, spheres, or ellipsoids of nearly constant shape; most other polymers are cross-linked to each other and so cannot independently change their shape.

EXTREME ELONGATION TAKING BROWNIAN MOVEMENT INTO ACCOUNT

The most widely accepted treatment of intrinsic viscosity is that of Simha: [16]

$$[\eta] = \frac{(b/a)^2}{15\left(\ln \dfrac{2b}{a} - \dfrac{3}{2}\right)} + \frac{3(b/a)^2}{15\left(\ln \dfrac{2b}{a} - \dfrac{1}{2}\right)} + \frac{14}{15}$$

where b is the major and a the minor axis of the particle considered as an elongated ellipsoid of revolution.

The point is that the hydrodynamic calculations of Jeffery completely ignored Brownian movement and, hence, assumed that the particle would lie in the plane of shear with its long axis in the stream line. Brownian movement will cause the particle to fluctuate from this position and thus get caught in the shear gradient and turn over end for end, thus producing an extra resistance to flow. There is no general solution for cases where shear and Brownian movement are comparable. The Simha equation given above is for the extreme limiting case that the shear is negligible in comparison with the Brownian movement. That is, that the particle turns and rotates in all directions regardless of the shear gradient, just as if the medium were at rest. This affords the maximum obstruction of which the particle is capable.

These two extreme cases are portrayed in Figure 10·6. The Jeffery formula is curve I, beginning with the Einstein coefficient, $a_1 = 2.5$ for a sphere. Likewise, Simha's curve II for the extreme possible influence of Brownian movement begins with the Einstein $a_1 = 2.5$ for a sphere.

In the latter extreme case the axial ratio b/a would have to be over 60 to produce a coefficient 100 times greater than that of Einstein. Not only this, but the particle would have to remain stiffly extended as a rod, which is found not to be the case. Thus, for polystyrene of molecular weight 270,000, Simha from the above formula deduces a thickness of 23.4 Å and a length of 1427 Å, whereas the fully extended molecule would have been 6490 Å. Thus, the polystyrene is curled up and flexible to some extent. "In different solvents, a given material may show a variation of 1 to 4 in intrinsic viscosity. This must be due to differences in the individual particles, and probably represents the effect of different solvate layers on the degree and compactness of coiling of extended molecules." [17]

Of course, the above discussion still refers only to extremely dilute solutions where the particles are assumed wholly independent. For the less dilute solutions, 3% by weight, of nitrocotton cited earlier in which the Einstein prediction was $\eta_r = 1.036$, if the Einstein coefficient were increased 100-fold, η_r would still be only 4.6; whereas the observed η_r is 1000.

FIGURE 10-6 Viscosity coefficient a, as function of axial ratio, f. I — Without Brownian movement (viscosity scale multiplied by ten). Ia represents portion of curve I below $f = 10$, with viscosity scale multiplied by 100 and axial ratio scale multiplied by 10. II — With Brownian movement. IIa represents portion of curve II below $f = 10$, both scales multiplied by 10.

Hence, the effect of extreme elongation is large, far larger than either of the two formerly advocated explanations. Nevertheless, it is still far smaller than the viscosity observed in real concentrations where the viscosity increases enormously with concentration (Chap. 20).

EFFECT OF RAMIFYING AGGREGATION OF PARTICLES: STRUCTURAL VISCOSITY

The most effective way of creating a large apparent viscosity with a small amount of suspended material is for the particles to link up to large loose ramifying aggregates extending through and effectively immobilizing much of the solvent. It is generally admitted that this is the explanation of the viscosity and other properties of soap solutions

FIGURE 10-7 Ferric oxide in amyl acetate, unprotected (left) and protected (right) by cellulose nitrate.

whose molecular weight is known and small. It is likewise generally assumed for Staudinger's less dilute solutions which he [18] therefore calls gel solutions. Many authors [19] have arrived at this conclusion, namely, that the apparent viscosity of most colloidal sols is really a structural viscosity of this kind. Electromicrographs of a ramifying aggregate of gold particles were reproduced in Figure 6·13.

First we may cite striking visual evidence that some such structure is present in a number of sols, such as cadmium sulfide, copper ferrocyanide, many dyes, starch, white of egg, and ammonium oleate. If a swirling motion is imparted to N/20 ammonia solution and a few drops of oleic acid are allowed to dissolve, the resulting ammonium oleate solution is only 8% more viscous than water, but it exhibits the remarkable property of not merely coming to rest like an ordinary liquid but of then reversing its direction of rotation, sometimes by almost a complete revolution.[20] No alternative has been found to the suggestion that this is evidence for the elastic distortion of ramifying aggregates within the solution. Such behavior is indeed a specific demonstration of their presence.

The possibility of pulling out a viscid liquid into threads, "spinning the liquid," would seem often to depend upon the presence of loose aggregates which can be distorted into elongated shapes without being broken up. Such liquid threads are often elastic.

McDowell and Usher [21] have published photographs, such as that in Figure 10·7, for five materials whose particles are shown to form ramifying aggregates when no protective colloid is present, but whose particles are separate and independent when suitably protected. Meanwhile, as shown in Figure 10·8, the viscosity of the suspension of separate protected particles is low and independent of the rate of shear, whereas that of the ramifying aggregates is ever so much greater and diminishes with the rate of shear as the aggregates are straightened or torn asunder. They conclude that "if rigid particles suspended in a liquid in which they are insoluble are not prevented from cohering — whether by an electric charge or by an envelope of a soluble substance — they will in time form aggregates, the presence of which will always cause the viscosity to be a function of the rate of shear; and which, if they are completely interlinked, will also impart rigidity to the suspension as a whole."

In accordance with this conception the author and his collaborators found that the best solvent for a cellulose derivative is the one that dismembers it most thoroughly and produces the lowest viscosity, in spite of the minor opposing factor that the best solvent is the one

FIGURE 10·8 Viscosity of mixture of amyl acetate and tetrachloroethane (△); same mixture plus 0.05 per cent rubber (□); 6 per cent suspension of starch unprotected (○); 6 per cent suspension of starch protected with 0.05 per cent rubber (×). Taken from McDowell and Usher, Text Ref. No. 21.
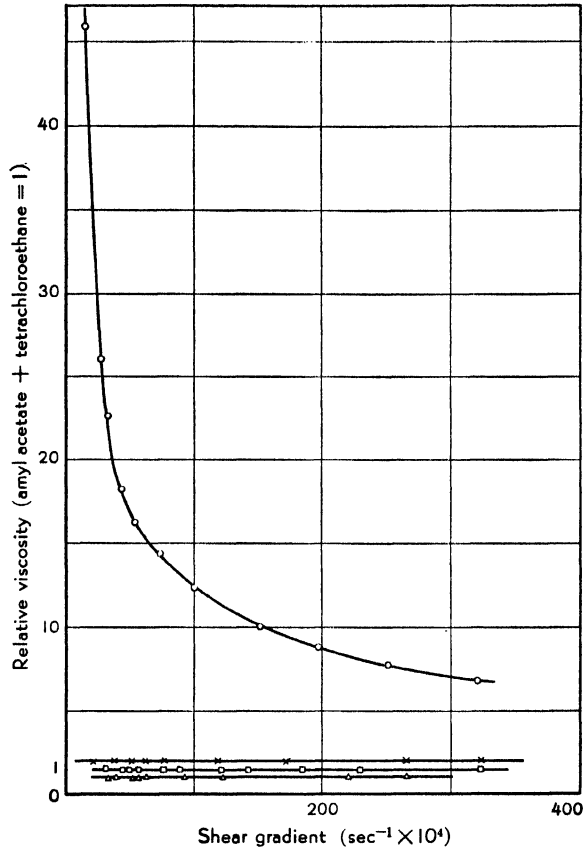
causing the greatest solvation. This is similar to the action of oil in releasing a glass stopper that is stuck in the neck of a bottle.

Structural viscosity of the type just referred to is the most effective method of minimizing the influence of temperature upon the flow properties of a liquid.

Flow of Pure Newtonian Liquids

The opportunity may be taken to refer to a general equation first derived by de Guzman in 1913 and now known as the Andrade equation [22] for relating the true viscosity of pure liquids to the temperature in such fashion that the logarithm of the viscosity is a straight line when plotted against the absolute temperature. From the empirical constants of the linear equation information may be derived regarding

the degree of association and its dependence upon temperature and molecular structure. Nissan and Clark [23] have extended this treatment to show that all the members of a homologous series have viscosities falling on one typical curve related to the absolute boiling point and to the viscosity at that temperature. Moreover, on their general diagram the liquid metals are separated from associated liquids by a narrow region containing the unassociated liquids or those that do not change their degree of association with temperature.

Types of Flow

It is useful now to plot diagrammatically four chief types of flow commonly observed, as in Figure 10·9. A fifth type will be discussed under the heading "Dilatancy."

Only the straight line A represents true viscosity with the constant η, such as is exhibited by gases and all liquid solvents, glass, pitch, and some colloidal sols where the flow is proportional to the force F. In general, however, Newtonian viscosity does not obtain and the study of the flow of materials is given the general name rheology.
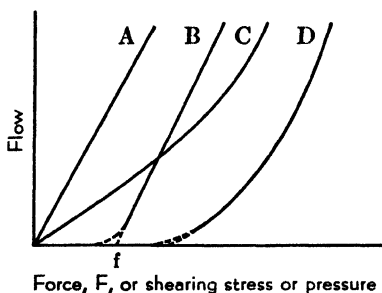
FIGURE 10·9 Types of flow: A, Newtonian; B, plastic with yield value; C, thixotropic; and D, thixotropic with yield value.

Curve B appears as a nearly straight line which, instead of pointing toward the origin, points, or may be so extrapolated, to a finite value of shearing stress termed the yield value f, the flow being proportional to $F - f$. If such a yield value exists, the flow is said to be "plastic." This type of flow is shown by butter, clay, certain paints, and asphalts. However, in practice some flow may occur below the yield value which is explained as being due to seepage of solvent through the mass of particles and sometimes to slippage against the walls of the vessel, so that the line B can be extended by a rather indefinite curve toward the origin. In an experimental confirmation [24] of the Reiner equation of flow relating plasticity and viscosity in a rotational viscometer of adjustable speeds, slippage was reduced or eliminated by cutting grooves in both concentric cylinders.

The more usual behavior of the more viscous colloids is shown by curves C and D which are curved throughout but differ in that D shows a more or less definite yield value; as with the transparent liquid

crystalline forms of soap solutions, such solutions are plastic. Curve C would be exemplified by cellulose derivatives, rubber, starch paste, chewing gum, and soap at not too low a temperature.

The diagrams are merely approximate and do not describe all the anomalies of flow that are met with in individual cases. It is evident that in many cases there is no way of inducing a uniform rate of shear and that flow may occur as a semirigid plug once the movement is inaugurated. In a tube the anomalies are usually most pronounced at the walls and the center. Plastic flow always results when solid particles touch each other, as in a concentrated mud or paint. If the flow is truly plastic or if there is a yield value, the material can be molded.

It is often useful to compare in a type C curve, as for example in a paint, the apparent viscosity coefficient, η_o, for very low shearing stress with the apparent viscosity coefficient, η_∞, for very high shearing stress where the aggregates are effectively dismembered. Williamson [25] defines the ratio $(\eta_o - \eta_\infty)/\eta_\infty$ as the "false body" which is a measure of structural effect. This shows why a paint can be designed to brush or spray or roll readily and yet not flow appreciably when applied to a vertical surface, η_∞ representing the former properties and η_o the properties after application.

Syrup will flow in a continuous stream from a wide tube without forming drops because its apparent viscosity is not changing greatly in this region of stress. A clay paste, on the other hand, falls out in blobs because of the great change in the neighborhood of the yield value. Soft soap is intermediate, and the blobs do not completely detach themselves from each other. Similarly, shortness in pastry or flour dough is associated with a particularly rapid decrease in viscosity with increase in shearing stress.[26]

THIXOTROPY

We have seen that when the apparent viscosity depends upon the rate of shear, as in curves C and D of Figure 10·9, this is due to loose linkages between many of the colloidal particles, resulting in loose packing and ramifying aggregates which are progressively broken down by increasing rates of shear.

It will be seen that the formation of ramifying aggregates giving an apparently high viscosity may proceed to such an extent that the system begins to resemble an elastic solid; a sol becomes a jelly. In these systems the effective, or structural, viscosity depends also upon the previous history and treatment of the system and how long it has

had to build up its structure undisturbed. When it is stirred or caused to flow, the structure is partially destroyed but slowly heals on standing. In 1927 Peterfi gave this familiar phenomenon the name "thixotropy," [27] that property of a body by virtue of which its consistency is temporarily reduced by previous deformation. In an extreme case the system may appear to be a solid jelly and yet upon shaking appear as a limpid fluid. But immediately after shaking it solidifies again, and a tube containing it may be inverted without flow occurring. The flow of such systems would be represented by curve C or D of Figure 10·9; or, rather, by a series of such curves of plastic flow with yield values displaced farther and farther to the right as time and opportunity are allowed for further building up of the structure.

Hauser and Reed [28] have observed the clustering of primary particles of bentonite and the linking together of the loose clusters, leaving open pores and channels within which individual particles still carried on Brownian movement. With the finest particles of bentonite 0.01% produces a structure and 0.85% a soft transparent jelly.

Practical illustrations are to be found not only in paints but in quicksand, possibly in muscle, and particularly in drilling muds. The latter, or mud-fluid, is now an important factor in petroleum oil field development. These thixotropic muds have been in use for more than sixty years.[29] They serve to bring up the rock chips during boring. Their density is controlled by the proportion of solid and liquid and by the addition of heavy materials, such as barium sulfate. Their effective viscosity is minimized by the addition of sodium tannate and of some of the pyro derivatives of phosphoric acid.

A numerical designation of the degree of thixotropy is sought by Goodeve and Whitfield [30] in the slope of the apparent viscosity when plotted against the reciprocal of the rate of shear, as the shear approaches a high value. This is taken to represent the rate at which the structure is being rebuilt while it is being torn asunder into its primary particles by the shear. The viscosity of a Newtonian liquid would result in a horizontal line.

In some cases a loosely packed structure is consolidated by gently rolling the vessel between the hands. However, a firm, white, solid curd, 0.1 or 0.05 N with respect to sodium palmitate, soon becomes a soft, wobbly mush exuding much of the formerly enmeshed liquid. Miss Worden and Miss McDowell in the author's laboratory utilized this unnamed property in studying sorption by fibers of soap curd. On the other hand, 0.1 or 0.2 N sodium laurate can be kept as sols for prolonged periods at room temperature, but shortly after jarring or slow

stirring they set to solid curds.[31] This process of accelerating the building up of structure by gentle treatment has since been given the name of "rheopexy" by Freundlich, but it is better known as "strain hardening."

DILATANCY

Dilatancy is a term introduced by Osborne Reynolds [32] in discussing the familiar fact that when one treads upon a wet, sandy beach, the sand appears dry around the foot until the pressure is released, when it becomes disproportionately wet. It is a result of the indifference of the particles to each other and their consequent tendency under the influence of gravity to lie or pack rather closely together. The liquid just fills the interstices. The thrust of the foot displaces the particles into a looser arrangement with larger interstices to take up liquid, but when left to themselves they settle close to each other again, leaving the excess water above.

This process has been imitated by Freundlich [33] with quartz powder having an average particle size of 1 to 5 microns containing 44 volumes of powder to 56 volumes of pure water. It has long been familiar in wet pastes of corn, rice, or potato starch with water. On standing, such systems leave an excess of supernatant liquid above a hard coherent clod or compact sediment. When mixed with the excess of water, they may be easily stirred, provided the stirring is slow, and they flow from a rod or spatula like a treacly liquid. They thus behave as a moderately

FIGURE 10·10 Newtonian flow, curve 1, contrasted with dilatant flow, curve 2.

viscous liquid, as is shown in Figure 10·10, where the straight line, number 1, represents the linear increase in flow with increase in pressure for a truly viscous liquid, and the first part A of curve 2 also rising linearly from the origin represents the similar behavior at low rates of stirring.

Now, however, if the attempt is made to stir rapidly, the system resists so vigorously that the rod may be broken. Or if the rod is dropped vertically from a height of several feet, it scarcely penetrates. The system cracks and breaks behind the moving stirring rod, and the actual flow does not exceed a certain maximum, independent of the

force. This is indicated by the horizontal portion, *B*, of the second curve in Figure 10·10 where the force is very greatly increased without effect upon the rate of flow. This is evidently due to the resistance caused by the particles jamming in irregular open structure in front of the moving rod instead of slipping smoothly past each other. The behavior is somewhat similar to pitch, which is brittle under shock and flows slowly but steadily under low shearing stress. Many synthetic plastic materials exhibit a similar behavior. In extreme cases the plastic may be as flexible as a sheet of rubber, but when hit sharply or snapped with the fingernail it breaks off in splinters.

Quartz and starch in other liquids, such as carbon tetrachloride or benzene, when allowed to settle occupy a much larger volume. They are loosely packed, due to linkage between particles, and are thixotropic. The sedimentation volume of cornstarch in water is reduced by only a few per cents (from 2.7 to 2.6) when a centrifugal field is applied, but in carbon tetrachloride it is reduced to a third (from 7.7 to 2.6).[34] Here there is a yield value in the loosely packed structure which resists sedimentation under gravity alone but gives under a larger force. The small addition of a surface active polar substance decreases the sedimentation volume greatly. Many studies of loose packing, as exhibited by large sedimentation volumes, conditioned by added electrolytes and other substances, have been made by von Buzägh.[35] This significance has not yet been fully evaluated or appreciated.

REFERENCES

Viscosity, Thixotropy, and Plasticity and Dilatancy

ANDRADE, Z. N. DA C., *Viscosity and Plasticity*, W. Heffer & Sons, Ltd., Cambridge, England (1947)

BARR, GUY, *A Monograph of Viscometry*, Oxford University Press, England (1931)

BINGHAM, E., *Fluidity and Plasticity*, McGraw-Hill Book Co., New York (1922)

BLAIR, G. W. SCOTT, *An Introduction to Industrial Rheology*, P. Blakiston's Son & Co., Inc., Philadelphia (1938)

British Standard Method for the Determination of Viscosity in Absolute Units, British Engineering Standards Association, 28 Victoria St., London (1929)

FREUNDLICH, H., *Thixotropy*, Actualités Scientifiques et Industrielles, No. 267, Hermann et Cie, Paris (1935)

HATSCHEK, E., *The Viscosity of Liquids*, G. Bell & Sons, Ltd., London (1928)

HOUWINK, R., *Elasticity, Plasticity and the Structure of Matter*, Cambridge University Press, England (1937)

NADAI, A., *Plasticity*, McGraw-Hill Book Co., New York (1932)

NOORD, N. V., First Report on *Viscosity and Plasticity*, Hollandische Uitgevers-Maatschappij (1935); Second Report (1938)

Text References

1. *Determination of Viscosity in Absolute Units*, British Engineering Standards Association, 28 Victoria St., London S.W. 1 (1929)
2. MCBAIN, J. W. and WATTS, O., *J. Rheology*, **3**, 437 (1932)
3. RICHARDSON, G. E., *Physics*, **6**, 273 (1935)
4. KING, A. M., *J. Soc. Chem. Ind.*, **41**, 147T (1922)
5. Discussed from a different standpoint by M. L. Huggins in "Cellulose and Cellulose Derivatives," *High Polymers*, Interscience Publishers, Inc., New York (1943), Vol. V, p. 943
6. EINSTEIN, A., *Ann. Physik*, **19**(4), 289 (1906); **34**(4), 591 (1911); *Z. Elektrochem.*, **14**, 235 (1908)
7. An exception is paraffin sol in carbon tetrachloride, where the constant is too small; as shown by K. H. MEYER and A. VAN DER WYK, *Helv. Chim. Acta*, **18**, 1067 (1935); *Kolloid-Z.*, **76**, 278 (1936)
8. HATSCHEK, E., "General Discussion on Colloids and Their Viscosities," *Trans. Farad. Soc.*, **9**, 10 (1913)
9. VON SMOLUCHOWSKI, M., *Kolloid-Z.*, **18**, 190 (1916); see also SAKURADA, I. and NAKASHIMA, T., *Kolloid-Z.*, **66**, 62 (1934)
10. MCBAIN, J. W. and SCOTT, D. A., *Ind. Eng. Chem.*, **28**, 472 (1936)
11. MCBAIN, J. W. and MCBAIN, M. E. L., *J. Am. Chem. Soc.*, **59**, 342 (1937); cf. EIRICH, F. and MARK, H., *Der Papierfabrikant*, **27**, 251 (1937)
12. JEFFERY, G. B., *Proc. Roy. Soc.*, **102A**, 163 (1923)
13. DOBRY, A., *Kolloid-Z.*, **81**, 190 (1937); see also Ostwald, Wo., *ibid.* 195
14. MCBAIN and SCOTT, *op. cit.*
15. KUHN, W. and KUHN, H., *J. Chem. Phys.*, **16**, 838 (1948)
16. SIMHA, R., *J. Phys. Chem.*, **44**, 25 (1940); *J. Applied Phys.*, **13**, 147 (1942)
17. KRAEMER, E. O., *Rheology Bull.*, **13**(3), 14 (1942)
18. STAUDINGER, H., *Chem. Zt.*, **59**, 733, 756 (1935)
19. For example, MCBAIN, J. W., SMITH, L. E. and HARVEY, C. F., *J. Phys. Chem.*, **30**, 312 (1926); MCBAIN, J. W., *ibid.*, 239
20. HATSCHEK, E., *Kolloid-Z.*, **37**, 25 (1925)
21. MCDOWELL, C. M. and USHER, F. L., *Proc. Roy. Soc.*, **131A**, 564 (1931)
22. SHEPPARD, E. C. and HOUCK, R. C., *J. Rheology*, **1**, 349 (1930)
23. NISSAN, A. H. and CLARK, L. V. W., *Nature*, **143**, 722 (1939)
24. GREEN, H., *Ind. Eng. Chem.* (Anal.), **14**, 576 (1942)
25. WILLIAMSON, R. V., *Ind. Eng. Chem.*, **21**, 1108, 1111 (1929); see also DE WAELE, A. and DINNIS, G., *Physics*, **7**, 426 (1936)
26. SCHOFIELD, R. K. and BLAIR, G. W. SCOTT, *Nature*, **136**, 147 (1935)
27. FREUNDLICH, H., "Thixotropy," *Actualités Scientifiques et Industrielles*, No. 267, Hermann et Cie, Paris (1935)
28. HAUSER, E. A. and REED, C. E., *J. Phys. Chem.*, **41**, 911 (1937)
29. EVANS, P. and REID, *Trans. Min. and Geol. Inst. India*, **32**, Dec. 1936, a complete monograph. AMBROSE, H. A. and LOOMIS, A. G., *Physics*, **1**, 129 (1932); **2**, 365 (1932); **4**, 265 (1933)
30. GOODEVE, C. F. and WHITFIELD, G. W., *Trans. Farad. Soc.*, **34**, 513 (1938); ARNOLD, J. E. and GOODEVE, C. F., *J. Phys. Chem.*, **44**, 652 (1940)
31. MCBAIN, J. W., CORNISH, E. C. V. and BOWDEN, R. C., *Trans. Chem. Soc.* (London), **101**, 2054 (1912)

32. REYNOLDS, OSBORNE, *Phil. Mag.*, **20**(5), 469 (1885); *Nature*, **33**, 429 (1886)

33. FREUNDLICH, H. and RODER, H. L., *Trans. Farad. Soc.*, **34**, 308 (1938); FREUND-
    LICH, H., *J. Phys. Chem.*, **41**, 901 (1937)

34. WILLIAMSON, R. V. and HECKERT, W. W., *Ind. Eng. Chem.*, **23**, 667 (1931); WILLIAM-
    SON, R. V., *J. Phys. Chem.*, **35**, 354 (1931)

35. VON BUZÄGH, A., *Kolloidik*, Theodor Steinkopff, Leipzig (1936); *Colloid Systems*,
    translated by O. and B. Darbishire, Technical Press, Ltd., London (1937); *Kolloid-
    Z.*, **47**, 370 (1929); **51**, 105, 230 (1930); **52**, 46 (1930); *Naturwissen.*, **18**, 444 (1930)

JELLIES AND GELS

The term gel is an omnibus word coined by Graham to include all the more solid forms of colloids, such as jellies, coagula, membranes, and natural and synthetic textile materials. The author draws a distinction between jellies and gels (cf. Chap. 19). The latter are related to the solid crystalline forms, as would be the case in a paste or a dried paste, whereas the jelly is a special form of sol, a colloidal solution. The emphasis in this chapter will be upon jellies [1] and those gels which are most similar to jellies. The reader should be warned that these terms have been and still are very loosely used by various authors. Some even restrict the term gel to true jellies and call the others pseudogels.

Structure of Jellies

When a colloidal sol contains particles which tend to join when they touch in favorable positions, forming ramifying aggregates, a high apparent or structural viscosity results. When this process develops further, so that the whole liquid begins to be effectively immobilized, the sol by progressive stages becomes a jelly without any line of demarcation between the two. In a jelly the unit is the primary particle, but these are sufficiently linked so that the whole liquid is enmeshed in the loose framework formed by aggregation. Much of the aggregation appears to be more or less linear, as if the particles were polar, or at least as if it were easier for a particle to add on to one other particle rather than directly to several at a time. Branching also freely occurs. The whole forms what has effectively been termed a "brush-heap" structure.

It is clear that such a jelly must be elastic and rigid and yet show yielding or relaxation under persistent stress. A very small amount of

solid material may suffice to form a jelly, a small fraction of a per cent of solid particles often sufficing to produce a rigid structure. As the concentration increases, the rigidity and hardness may become that of a glass. Indeed, glasses may be classified as concentrated jellies. Figure 11·1 indicates very diagrammatically the supposed structure.[2] It may be compared with the "distinctly dendritic" structure of Figure 10·7 and with similar photographs which Usher [3] obtained from other greatly diluted gel- or jelly-forming systems.
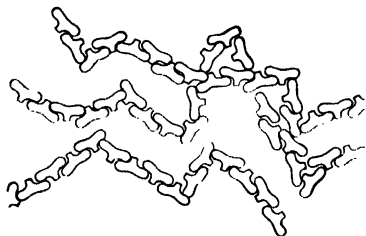
FIGURE 11·1 Diagrammatic representation of a jelly.

It is thus seen that for the formation of a jelly two properties are necessary; first, sufficient stability or solvation, and secondly, the power of joining at points of contact.

It is important to note that all jellies may be prepared so as to have no particles visible in the ultramicroscope and yet retain their typical mechanical properties. The ultramicroscope fails to resolve the small, solvated primary particles and therefore does not show the ramifications of their loose aggregates. Light scattering likewise corresponds to the small primary particles which in a ferric hydroxide sol may be only of the order of 20 Ångström units, even when aggregation has gone so far that the ferric hydroxide does not pass through ordinary filter paper.

The structure indicated enables us to predict correctly many of the chief properties of such systems. For example, the mechanical properties are not those of an emulsion, and the elasticity increases much more rapidly than the concentration, generally to a second or higher power. The arrangement of the structural aggregates will depend upon how the system was moving when the structure set. Since, under stress, the solvent will slowly move past the myriad enmeshing particles, there will be a slow relaxation which will leave the structure permanently deformed. If double refraction has been produced by a strain, this will remain after the structure has relaxed into its new position, and there will be little or no tendency to revert to the original form, except in very concentrated systems. On drying, such a deformed specimen, if it shrinks progressively, will deform still further.

When bubbles are formed in an ordinary liquid, they are spherical; but Hatschek found that when they are generated in a jelly they have the shape of lenticular capsules, even when the jelly has appeared to

be optically isotropic. The compressibility of a jelly is very nearly that of a liquid. Thus if a piece of rubber is stretched in one direction, its volume is practically unaltered, the cross-sectional area being correspondingly diminished. If a jelly is stuck to the walls of a container, its structure will be evidenced by the way in which some of the strain will be borne by the walls; and in this way, when compressed by a column of mercury, droplets may shoot into the center of the jelly. To study the mechanical properties of the jelly, the jelly should be motionless and formed loose from the walls, and a whole lump of it should be investigated.

Poole [4] has given quantitative consideration to the behavior expected from a viscous liquid containing a random arrangement of elastic fibriles. He takes the simple Maxwellian model of a loose-fitting piston in a cylinder filled with a viscous liquid, the submerged piston being attached with a helical spring to the bottom of the cylinder and connected to the exterior by another helical spring. The result of a pull on the latter depends upon the time and is compounded of viscous flow of the liquid and of elastic flexure of the fibriles. The results agreed with the experimental behavior of Young's modulus of elasticity of jellies of cellulose acetate in benzyl alcohol.

Two other properties which might be deduced from this structure refer to the effect of gelatinization upon diffusion and upon electrical conductivity. Dilute jellies affect these properties very little, movement being almost as free as in pure solvent. Concentrated jellies, on the other hand, because of their crowded particles greatly diminish conductivity and stop the diffusion of particles that are too coarse for the interstices, as was pointed out in the introductory chapter and in discussing ultrafiltration.

The subject of Liesegang rings will be briefly referred to in the chapter on diffusion.

Types

Jellies and gels are formed in a variety of ways, and they have not been strictly classified. A first or heat-reversible type of true jelly is the most familiar; one example is edible table jelly made by dissolving gelatin in warm water and cooling it until it sets. Other examples are agar, or sodium oleate, or pectin jellies, or soluble methyl cellulose in water, or cellulose acetate in benzyl alcohol. The solvent is always selective or specific. In most cases a clear, transparent jelly appears to liquefy or melt on heating over a rather narrow range of temperature

which is below 33° C for aqueous gelatin, 23° C for sodium oleate, and 93° C for agar. Methyl cellulose in water and cellulose nitrate in alcohol show the unusual inverse effect of liquefying on cooling. In all cases there is a good deal of hysteresis, that is, a time lag between the cooling and the setting to be expected of the jelly. Thus a sol may persist for a time at temperatures well below the setting point while the particles are gradually linking. Although usually setting occurs on cooling, the greater the undercooling the slower the setting, since the *rate* of setting has a high positive temperature coefficient. It may be mentioned that a sol of ferric oxide containing sodium chloride sets at a rate which increases 4-fold for a rise of temperature of 10° C. The rate of setting of jellies and gels may be represented by an empirical equation [5] connecting the setting time t with concentration c and absolute temperature T and an activation energy Q in the form

$$t = ac^{-Q/RT}$$

where a is a constant.

The setting point, or rather zone, is greatly dependent upon the presence of other substances. For gelatin setting is favored by sulfates and by such hydroxyl compounds as sucrose and glycerin. It is hindered by chlorides and nitrates, and it is actually prevented by potassium iodide, thiocyanate, and urea. The elasticity is similarly affected. The urea might be considered to react with the terminal amino groups, but these substances have the opposite effect in ferric hydroxide sols. Such influences of electrolytes were discussed in Chapter 9, devoted to Lyotropic Series.

A second type of jelly is the irreversible type. For example, a sol of silicic acid or alumina sets irreversibly, merely upon keeping. The rigidity steadily increases with time. These jellies are noted for the fact that when they have been dried below a certain content of water they are rigid, porous masses which will now imbibe any wetting liquid whatsoever. This indifference to the nature of the liquid filling the structure is in sharp contrast to the previous type where the liquid has always to be a solvent capable of forming a sol and where no such empty pores arise on drying, because the jelly shrinks progressively as the solvent evaporates. Holmes [6] has prepared a series of so-called singing gels of this rigid, irreversible type by allowing them to form with suitable concentrations in a series of test tubes. When tapped they vibrate audibly.[7]

A third type is the unstable jelly that is on its way, not to coagulation, but to crystallization or recrystallization. Many examples are

to be found among the alkaloids and the urates but the best known are Hardy's [8] azomethine and Hatschek's [9] camphoryl phenylthio-semicarbazide. These crystals dissolve in a number of hot, organic solvents to form clear solutions which, on cooling, set to clear true jellies. Some can be kept for less than a minute, but others for months before the system degenerates into a mass of crystals lying at the bottom of the liquid solvent. However, Hatschek's toluene jelly may actually be stable when it contains 2% or less of solid, for specimens were kept apparently unchanged for months. Any excess above 2% crystallizes out within the clear jelly, which shows a melting tempera-ture of about 50°, the crystals not all dissolving until much higher temperatures. Some of the soap jellies such as sodium stearate and aluminum palmitate in toluene or xylene, investigated by von Weimarn and by the author, are actually stable.[10] Other jellies which tend to crystallize are produced by reactions forming insoluble substances such as manganous arsenate.

A fourth type of jelly or gel is represented by the ultrafilter mem-branes already discussed. Elford's membranes of pore size 100 Å for example are perfectly transparent, although those whose pore size is of the order of the wave length of light, such as 8000 Å, are necessarily white and opaque like the petals of a lily, for the same reason in both cases. Elford has shown that the desired structure is comprised of small colloidal micelles.

Thus all these types are comprised of small colloidal particles and differ chiefly in the concentration, in the strength of the linkages between particles in contact, and in the deformability of the branching structure as to how strongly it resists closer packing, avails itself of opportunity of a looser arrangement, or tends to shrink spontaneously to still closer packing. The latter phenomenon leads to exudation of a portion of the liquid and is called syneresis. If carried further it results in coagulation.

Syneresis [11] is observed with gels of sodium soaps if sufficiently dilute. The liquid extruded is merely a saturated solution. Various additives accelerate or retard syneresis.

There are two ways of preventing loss of solvent by syneresis. One is to add something that is sufficiently soluble and also has an affinity for the solid structure so as to produce osmotic pressure, which will keep the structure filling the available liquid. The other method of preventing syneresis is to add an insoluble, extremely finely divided mechanical framework, such as curd fibers of sodium stearate or palmi-tate, which will stiffen the structure so that it cannot shrink and then

have sufficient rigidity to inhibit syneresis. This method is due to Hattiangdi,[12] who prevented the syneresis of sodium oleate gels and jellies in pinene by adding a gel of sodium stearate fibers. Fischer [13] found that 1 gm of insoluble curd fibers of sodium arachidate could hold 82 gms of ethyl alcohol without syneresis.

Hydration and Dehydration of Jellies

A formerly held explanation of the solid nature of jellies was that a large part of the liquid was combined with the colloidal particles. This view has been abandoned because it is found that the solid materials sorb less than their own weight of nearly saturated vapors, the amounts diminishing with the poorness of the solvents.[14] Another test is to add a reference substance to the jelly and analyze the syneretic liquid exuded or squeezed out. If it is more concentrated than the jelly, some of the solvent must have been bound, or held, otherwise than as free solvent by the solid structure. In this way Gaunt and Usher [15] found that silica gel held only 1 H_2O to 1 SiO_2, whereas precipitates of silica, ferric oxide, and alumina held 1.5, 3, and 6 H_2O, respectively. Thus gelatinization does not necessitate high hydration. Similarly it is found [16] that most of the water in a jelly can be frozen, leaving for example [17] 0.5 gm water to 1 gm gelatin. However, this is not a rigorous method, for, on the one hand, isolated pockets of water may not freeze, and, on the other, the lowering of vapor pressure may itself displace the equilibrium.

A study of the process of dehydration of jellies, to which van Bemmelen especially devoted many years,[18] sheds light upon their structure. Van Bemmelen was impressed by the fact that the vapor pressure curves showed no sharp breaks even when dehydration was carried to dryness. However, his conclusion that therefore particles of the oxides do not form definite hydrates is invalidated by a consideration that even these must obey the universal laws of sorption for surfaces.

The best studied example is the jelly which forms when solutions of sodium silicate are acidified and allowed to stand. The electrolytes may be dialyzed out of the very soft jelly, whose original gross composition might be 1 SiO_2 with 50 to 100 H_2O. Therefore as the jelly is allowed to lose water by evaporation, most of it is given off nearly as readily as pure water, and it is only when a couple of moles of water are left that we enter upon the diagram reproduced in Figure 11·2. The jelly shrinks steadily until less than 1.5 moles of H_2O are left. Then it becomes rigid. Up to and including this distinctive point O,

addition of small amounts of water suffices to restore the full vapor pressure of water, the process being completely reversible along the curve $O\ O_2\ O_3$, as is represented by the arrows in the diagram for this, or the nearly parallel steep dotted curve arising from any chance point at which dehydration was interrupted and hydration resumed. Up to this point the jelly is transparent and tends to shrink spontaneously on standing by exuding or allowing to leak out a part of its water (syneresis).
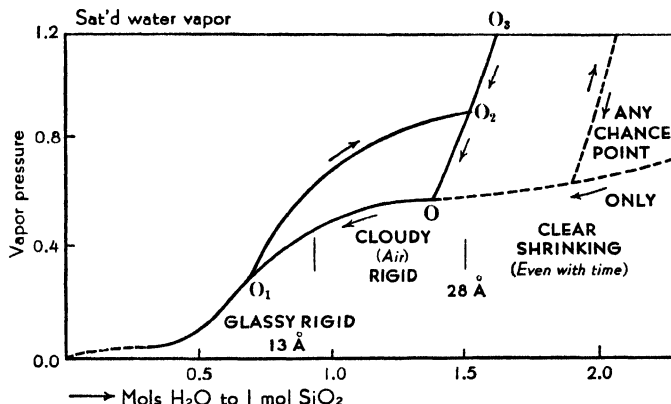
FIGURE 11·2 Dehydration of silicic acid jelly.

From O onwards the jelly becomes opaque or white and the process exhibits hysteresis, as is shown by the loop $O\ O_1\ O_2$ and the accompanying arrows. Beyond O_1 the usual reversible sorption curve is followed, the solid now being a clear glass. The cloudy interval is due to the liquid having been abstracted from patches of the loosely packed particles, leaving the interstices filled with vapor. The empty volume may be filled with any liquid such as water, alcohol, or benzene and behaves as the equivalent of capillaries of 28 Å effective radius at O. The vapor pressure falls at O_1 as if all the capillaries above 13 Å had been emptied.

Aerogels [19]

The interpretation just given is that the real structure of the jelly consists of ultramicroscopic particles which remain irregularly and loosely packed even when the jelly has shrunk to rigid form. The shrinking is evidently the result of the compressive effect of the surface tension of the evaporating liquid. Otherwise there is but little reason for the jelly to shrink.

The mechanism of the shrinking is as follows. As the solvent on the exterior evaporates, the surface first touches the high spots of any interior structure and then on further evaporation sinks in between them. The surface tension then must pull in any projecting parts of the structure, and the jelly itself will therefore shrink. The force involved follows from Cantor's law for the hydrostatic pressure due to a curved meniscus of radius r (between asperities), $p = 2\sigma/r$ where σ is the surface tension.

If this is found to hold it not only confirms the mechanism just indicated but demonstrates that the conception of a jelly as having a loose framework of colloidal particles is correct. This is very neatly proved by Kistler,[20] who adopted the device of replacing the water by any other liquid with a low critical temperature and warming the jelly in a completely enclosed space until the liquid had become vapor, then letting it escape quietly, all liquid surface having been eliminated. Here the jelly does not shrink and is extremely light, soft and compressible, highly opalescent and opaque by scattered light, but perfectly clear for directly transmitted light.

Aerogels have been made from most of the recognized jellies as well as from the oxides of many metals. They form good insulators; when powdered they mix rapidly and uniformly with pigments, and they are often excellent catalysts or supports for other catalysts.

Swelling [21]

An "elastic" jelly is commonly, although not very aptly, defined as one that when immersed in a suitable liquid swells either to an apparent limit or indefinitely to form a sol. A "rigid" jelly or gel is one that scarcely swells, like silica gel or charcoal. The limit of swelling depends upon the solvent, which must be specifically suitable. Gelatin or starch will swell in water and not in petroleum ether; rubber in ether or toluene but not in water; nitrocellulose in ketones, indefinitely, or to a limit in toluene or alcohol but not in water. An elastic jelly may be hardened to a rigid jelly, as for example, gelatin by formaldehydes or alcohol.

The swelling of gelatin in water is much affected by acids, bases, and electrolytes. Those that increase elasticity diminish the swelling and vice versa.

The rate of swelling is very great and may be practically instantaneous, as, for example, a thin sheet of gelatin exposed to moist air or rubber dipped in benzene. Its physiological importance is evident,

since muscular action is an example of the effect of change in concentration and pH of the liquids in which the fibrils are bathed.

Elastic jellies swell in a similar manner in their initial stages whether or not this proceeds to a limit. The extent often depends upon the previous history, such as intermediate drying, and time, or a previous or initial concentration. The structure of a jelly is clearly evidenced by Schroeder's paradox by which the upper half of a strip of jelly held in a saturated vapor swells to several times less than the thickness of the lower half which has been immersed in a liquid. The explanation is that in the liquid the mechanical strains are more fully released. Some materials such as wood or cellophane swell chiefly in one or two directions. Examination with x rays is a method of testing whether the swelling is only between the particles, leaving the radiogram unchanged, or within them, or both.

For elastic jellies whose particles pack more and more closely on dehydration without jamming, as in rigid jellies, one has a swelling diagram, without hysteresis, such as is shown in Figure 11·3. As the swelling proceeds, the particles undoubtedly sorb more liquid, but pore or capillary interstices are simultaneously developed which immediately fill with liquid.

FIGURE 11·3 The swelling and drying of elastic jellies.

This behavior is shown by gums, cellulose derivatives, casein, albumin, hemoglobin, vitellin, and many other materials.

The total heat evolved during complete swelling amounts to about 20 to 30 calories per gram of dried substances. However, the initial heat, for the first solvent taken up, is at the rate of hundreds of calories per gram, soon falling off toward an immeasurably small value. It so happens that the heat of swelling is very nearly equal to the free energy evolved, namely $RT \ln p/p_s$.

The swelling pressure is thermodynamically related to the relative vapor pressure, p/p_s, by the formula:

$$P = -\frac{1}{V} RT \ln p/p_s$$

where V is the volume of 1 mole of liquid solvent. Thus the initial swelling pressure is infinitely great and for $p/p_s = 1\%$ is of the order of 5000 atmospheres, falling to zero where $p/p_s = 100\%$. Old examples of the impressive magnitude of the pressures against which initial

swelling can occur are the use of dried peas thrust into cracks and then wetted when quarrying stone, the shortening of a rope carrying a heavy weight when water is thrown onto it, and the growth of a mushroom lifting a heavy paving stone.

As would be expected from their structure, jellies upon swelling get weaker and less elastic or rigid. They stretch like a liquid with practically constant volume. If, as in rubber, anisotropic particles form and become ordered by stretching, this is accompanied by the absorption of heat and, conversely, the stretched jelly contracts if warmed, as the crystalline patches disappear (Chap. 23).

Gels

Gels differ from jellies in that they bear a much closer relation to the crystalline structure of the solid which is responsible for forming the gel. In a gel one may expect x-ray examination to show that much of the original pattern of the pure solid is retained; whereas in a jelly this pattern is lost, or there has been a complete reorganization of the molecules to form the particles of the jelly and of the sol to which the jelly is so closely related. A crystalline solid such as aluminum dilaurate may be prepared in different physical conditions showing progressive swelling without losing the distinctive x-ray pattern or dissolving to form a sol. In such cases there is a sharp transformation from gel to jelly although it is possible for the gel to be in equilibrium with a jelly or a sol and a system may consist of a mixture of these.

<div align="center">REFERENCES</div>

Text References

1. HATSCHEK, E., *J. Soc. Chem. Ind.*, 48, 389 (1929)
2. MCBAIN, J. W., *J. Chem. Ed.*, 6, 2115 (1929)
3. USHER, F. L., *Proc. Roy. Soc.*, 125A, 150 (1929)
4. POOLE, H. J., *Trans. Farad. Soc.*, 21, 114 (1925); 22, 82 (1926)
5. HATTIANGDI, G. S., *J. Coll. Sci.*, 3, 207 (1948)
6. HOLMES, H. N., KAUFMANN, W. E. and NICHOLAS, H. O., *J. Am. Chem. Soc.*, 41, 1329 (1919)
7. KOHLRAUSCH, F., *Z. physik. Chem.*, 12, 773 (1893)
8. HARDY, W. B., *Proc. Roy. Soc.*, 87A, 29 (1912)
9. HATSCHEK, E., *Z. Chem. Ind. Kolloide*, 11, 158 (1913)
10. MCBAIN, J. W. and LAING, M. E., *Trans. Chem. Soc.* (London), 117, 1507 (1920); MCBAIN, J. W., *Colloid Symp. Monog.*, IV, 1 (1926); *Kolloid-Z.*, 40, 1 (1926); MCBAIN, J. W. and MCCLATCHIE, W., *J. Phys. Chem.*, 36, 2571 (1932); HATTIANGDI, G. S., *J. Sci. Ind. Research* (India), 4, 491 (1946)
11. PRASAD, M., HATTIANGDI, G. S. and MATHUR, K. N., *Proc. Indian Acad. Sci.*, 21, 105 (1945); PRASAD, M., HATTIANGDI, G. S. and ADARKAR, S. P., *ibid.*, 22, 320 (1945)

12. HATTIANGDI, G. S., *Proc. Indian Acad. Sci.*, **22A**, 320 (1945)
13. FISCHER, M., *Chem. Eng.*, **27**, 184 (1919)
14. For example, MCBAIN, J. W. and collaborators, *Trans. Farad. Soc.*, **29**, 1086 (1933)
15. GAUNT, H. and USHER, F. L., *Trans. Farad. Soc.*, **24**, 32 (1928)
16. LLOYD, D. JORDAN and MORAN, T., *Nature*, **132**, 515 (1933)
17. MORAN, T., *Proc. Roy. Soc.*, **112A**, 30 (1926); *ibid.*, **135A**, 411 (1932)
18. VAN BEMMELEN, J. M., *Die Absorption, Gesammelte Abhandlung*, Theodor Steinkopff, Dresden (1910)
19. KISTLER, S. S., *J. Phys. Chem.*, **36**, 52 (1932)
20. KISTLER, *loc. cit.*
21. KATZ, J. R., *Ergebnisse der exakten Naturwissenschaften*, **3**, 316 (1924); **4**, 154 (1925)

COAGULATION, PROTECTIVE ACTION,

AND SENSITIZATION

Unstable or irreversible colloids are those that are readily coagulated by slight contamination. However, even the most unstable colloid has some degree of stability, for otherwise it would not exist except for the briefest period; whereas if kept sealed and away from all contamination, it may last indefinitely. Generally, limited stability is due to the lack of sufficient or suitable stabilizing agent, because this class of colloid cannot supply its own stabilizing agent. In contrast, stable colloids that are self-stabilizing, such as soaps, detergents, proteins, and dyes can tolerate fairly high concentrations of many electrolytes and nonelectrolytes provided that there are no specific interactions. This subject is touched upon in Chapters 17 and 18, in addition to the brief mention in the final section of this chapter, "coacervation."

Seabra [1] has elaborated on Lumière's [2] hypothesis of irritation of the sympathetic nerve terminations of the vascular endothelium caused by incipient coagulation of body fluids, resulting in "colloidoclastic disturbance," allergy, anaphylaxis, and shock.

Before dealing with the important effects of electrolytes, those of nonelectrolytes may be briefly described.

NONELECTROLYTES WITH IRREVERSIBLE OR UNSTABLE COLLOIDS

The effect of nonelectrolytes upon irreversible sols is very simply stated. Usually there is no observable effect. For example, many organic substances soluble in water do not affect arsenic trisulfide.[3] Sometimes there is an influence which can be traced to the added sub-

stances, such as alcohol lowering the solubility of the stabilizing agent. An example of this would be ferric hydroxide prepared from sulfate where the stabilizing sulfate is soluble in the alcohol, whereas ferric hydroxide prepared from chloride would not be noticeably altered. Stable colloids are sometimes sensitive to specific reagents. Alcohol, for example, is a well-known coagulant of certain proteins. This will be discussed in connection with the respective colloids.

NONELECTROLYTES WITH SUSPENSIONS

Suspensions behave with respect to added nonelectrolytes with much the same indifference as do the irreversible sols. For example, thirteen organic substances in fairly large concentrations were tested with kaolin without producing coagulation. Nonelectrolytes may modify the influence of electrolytes on suspensions or unstable colloids; occasionally by stabilizing them, but usually by sensitizing them toward coagulation.

ELECTROLYTES WITH IRREVERSIBLE COLLOIDS

In a few cases an electrolyte supplies the deficient stabilizing agent or peptizes the colloid. For example, hydrogen sulfide stabilizes arsenic trisulfide and ammonia stabilizes gold sol. Sometimes, if the electrolyte is added slowly and carefully, the result is a transparent jelly, as with Weiser's chromium trioxide.

However, in the vast majority of cases coagulation is the result. This is illustrated on a massive scale by the formation of deltas when muddy rivers reach salt water.

For each electrolyte there is a small threshold concentration which has to be exceeded in order to produce noticeable coagulation. But the colloid is modified even below this threshold value, as is seen when red gold sols produce the more coarsely aggregated blue gold sols or when electrokinetic movement is diminished. The threshold value is not accurately fixed because the rate of addition and the rate of mixing the solutions are important. In the ultramicroscope the coagulation is seen to be gradual and continuous and fairly rapid, occurring in a few seconds or minutes.

Fairly broad regularities can be observed. Particular ions are effective, especially those of highest valency and of opposite sign of charge from the colloidal particle. Amongst univalent ions the hydrogen ion and those of high molecular weight, such as the alkaloids and dyes

are especially effective. In surveying the chemical data it is very clear that the valency of the ion of opposite charge from the colloidal particle is the most important single influence. But it is equally clear that the nature of this ion and the nature of the other ion of the added electrolyte cannot be left out of account. The regularity of the influence of valency was used by Galecki in 1908 to prove that beryllium is divalent.

Table 12 presents some typical data due to Freundlich, working under constant arbitrary conditions, and it shows the concentration of the added electrolyte in millimoles per liter required to produce comparable coagulation in a given time. The results are given for one typical negatively charged colloid, arsenic trisulfide, and for the typical positively charged colloid, ferric hydroxide.

TABLE 12

AMOUNT OF ELECTROLYTE IN MILLIMOLES PER LITER REQUIRED TO PRODUCE COAGULATION OF A TYPICAL NEGATIVE SOL AND A TYPICAL POSITIVE SOL

As$_2$S$_3$ (negative)		Fe(OH)$_3$ (positive)	
Electrolyte	Millimoles	Electrolyte	Millimoles
$\frac{1}{3}$ K$_3$ citrate	240	KCl	9.03
KAc	110	KNO$_3$	11.9
NaCl	51.0	$\frac{1}{2}$ BaCl$_2$	9.64
KNO$_3$	50.0		
KCl	49.5		
LiCl	58.4		
$\frac{1}{2}$ K$_2$SO$_4$	65.6		
NH$_4$Cl	42.3		
HCl	30.8		
MgCl$_2$	0.72	K$_2$SO$_4$	0.20
MgSO$_4$	0.81	MgSO$_4$	0.22
CaCl$_2$	0.65	K$_2$Cr$_2$O$_7$	0.19
ZnCl$_2$	0.69		
UO$_2$(NO$_3$)$_2$	0.64		
AlCl$_3$	0.0093	H$_3$ citrate	0.0007
Al(NO$_3$)$_3$	0.095		
Ce$_2$(SO$_4$)$_3$	0.092		
Morphine Cl	0.425		
Neufuchsin	0.114		

In the first place we find comparable behavior of cations of a given valency toward arsenic trisulfide and of anions of a given valency toward ferric hydroxide. Next we see that there is a wide range of effectiveness for a given valency. Finally, the other ion present has a minor effect.

Hatschek's students were required to memorize the following table in order to compare and contrast a few typical sols:

TABLE 13

COAGULATION OF COLLOIDS BY VARIOUS ELECTROLYTES

Sols	Concentration in Millimoles per Liter of Mixture					
	Charge	NaCl,	HCl	CaCl$_2$,	BaCl$_2$	AlCl$_3$
Prussian blue or As$_2$S$_3$	—	51		0.65		0.09
Au or Pt	—	2.5			0.06	0.01
Mastic	—	1000	120	25	22	0.2
Fe(OH)$_3$	+	9.2		0.02	K$_2$SO$_4$	—

Therefore, if 18 cc of As$_2$S$_3$ are to be coagulated by 2 cc of NaCl, etc., the latter must contain 10 × 51 millimoles NaCl per liter. (This would amount to $\dfrac{10 \times 51}{500} = 1$ millimole in the 2 cc.)

The amount of electrolyte required to produce coagulation depends upon the total surface exposed by the colloidal particles. Hence more concentrated sols, or more highly dispersed sols with finer colloidal particles, require distinctly more electrolyte for coagulation. This effect is most marked for univalent ions, much less so for ions of high valency.[4]

Various authors have attempted to give general numerical values for the ratios of concentrations required for mono-, di-, and trivalent ions. The effect was first observed by Schulze in 1883 and is usually referred to as the Schulze-Hardy rule. However, if an effective concentration of trivalent ion is taken as unity, the numerical values of divalent ions range between 7 and 200 times greater, and those of monovalent ions between 500 and 10,000 times greater.

Freundlich related the effectiveness of valency to the ease of sorption of the corresponding ions. Other authors, such as Zsigmondy, complain that these rules leave out the whole chemistry of colloids.

In discussing the mechanism of these effects we must remember that they are due to interaction of the electrolyte with the stabilizing agent of the colloid and that the added ions may modify or even replace the original stabilizing agent. This interaction is clearly demonstrated by the fact that the coagulum always contains definite amounts of the ion which was effective in producing the coagulation. Thus arsenic trisulfide coagulum retains some of the precipitating cation and liberates an equivalent amount of hydrogen ion in the form of acid in the free liquid. This liberation of an equivalent amount of hydrogen

ion has been verified for arsenic trisulfide and gold sols by Freundlich, and for sols of sulfur and of stearic acid by Bolam.[5]

The following table shows the amounts of cation retained by arsenic trisulfide, although the experiments were carried out in three different countries with unrelated specimens.

TABLE 14

THE MILLIEQUIVALENTS OF CATIONS CARRIED DOWN IN THE COAGULUM CONTAINING 110 MILLIMOLES OF As_2S_3

Cation	Milliequivalents	Investigator
Ba^{++}	2.11	Linder and Picton, 1895
K^+	2.01	Whitney and Ober, 1902
Ca^{++}	2.46	Whitney and Ober, 1902
Sr^{++}	2.02	Whitney and Ober, 1902
Ba^{++}	2.72	Whitney and Ober, 1902
$C_6H_5NH_3^+$	1.81	Freundlich, 1907
Cation of neufuchsin	1.86	Freundlich, 1907

The number of equivalents is evidently independent both of the valency of the coagulating ion and of the concentration required for coagulation. Treatment of these coagula with solutions of other electrolytes results in ion exchange, equivalent for equivalent.

More careful experiments have been carried out by Annetts and Newman[6] showing the amounts of cation and of anion present in various coagula from gold sols stabilized with different agents.

Between 90 and 99% of the precipitating ion is adsorbed, and 95 to 98% of the ion of like charge with the colloid remains in the supernatant liquid.

Usher[7] has carried out experiments distinguished by the fact that the area of the monodisperse suspension of gamboge particles was known. The majority of the coagulating ions were found to be "employed in displacing ions originally attached to the particles but not contributing to their effective charge. The total charge on the flocculating ions is therefore greatly in excess of the effective charge considered responsible for the stability and the electrokinetic behavior of the particles." He emphasizes that "among ions of the same valency the flocculating power is determined by the solubility product of the compound formed from the flocculating and stabilizing ions." He therefore adheres to the theory that the stabilizing agent on a colloidal particle is only partially dissociated.

The interplay of the effect of high valency of one ion and the related concentration of the other can result in the so-called irregular series.

Here the addition of excess of the high valent coagulating ion can reverse the charge of the colloid if the addition is rapidly made so that the concentration is built up too quickly to produce coagulation. This has been illustrated in Bredig's experiments with negatively charged colloidal metals to which ferric or aluminum chloride or basic dye has been added. If now still further amounts of the electrolyte are added the chloride ion reaches a concentration where it in turn coagulates the now positively charged colloid. Thus coagulation of the negative colloid occurs with moderate addition, no coagulation with a greater addition, and finally coagulation of the colloid of the reversed sign when a still more concentrated solution is added.

The theoretical aspects of coagulation by electrolytes to which various authors have drawn attention are neutralization of the electric charge on the colloid, reduction of electrokinetic movement below a critical value, the "compression of the diffuse electrical layer," the depression of the dissociation of the stabilizing agent, the formation of an insoluble compound with the stabilizing agent, the replacement of the stabilizing agent by ion exchange, the sorption of the active ion, and the relation of coagulation to the activity coefficient of the coagulating ion. As regards the actual process of coagulation, this necessarily involves collision of the primary particles so that they may stick together if they are unstable. The kinetics of this were formulated by Smoluchowski in 1912 on the alternative assumptions that every collision is inelastic or that only a certain fraction is effective, but there are such wide departures from his formulations that they are not quoted here.[8]

Coagula, flocculations, and gelatinous precipitates differ from the corresponding massive crystalline substances in density, color, increased solubility, etc. In many the particles still persist, though more or less stuck together. Sometimes they may be brought back into colloidal solution by suitable reagents or ion exchange. This will be discussed in connection with coacervation.

COLLOID WITH COLLOID

There are three main kinds of interaction between colloid and colloid: (1) They may mutually precipitate (or sometimes form coacervates) if in proper proportions; (2) One colloid may make another easier to precipitate by electrolytes; this is called sensitization; (3) One may protect or stabilize the other; this is called protective action.

Coagulation of One Colloid by Another

Usually for mutual coagulation two colloids should be of opposite sign. Thus arsenic trisulfide and gold sol, both negative, show no visible reaction.

The mutual coagulation of two colloids was a striking phenomenon first studied by Graham, then by Linder and Picton, and still later by others such as Biltz. The standard example is the interaction of arsenic trisulfide and ferric hydroxide. The chief characteristics for the effects of quickly mixing sols with each other are shown diagrammatically in Figure 12·1. The effect of initial additions is slight but rapidly increases to quantitative coagulation of both colloids, generally somewhere near the isoelectric point, where the two colloids of opposite charge have just neutralized each other. The coagulum is a solid, not a liquid coacervate. From the approximately symmetrical diagram it is clear that a greater initial addition of colloid may result in an oppositely charged sol without producing coagulation. The chief properties of the mixture are those of the colloid present in excess. However, it should be pointed out that gradual addition of one sol to the other will finally result in complete coagulation of the second sol and that further addition of the first sol will not bring back the coagulum into solution except where the added sol is a stable protective colloid. (See the discussion under that heading below.)

FIGURE 12·1 Diagram of mutual coagulation.

The interaction is far too rapid and too dependent upon the speed of mixing for true equilibria or for the formation of true compounds. The adherence of the oppositely charged particles has been observed in the ultramicroscope.

The obvious explanation is that advanced by Billitzer in 1905 — mere electrical equivalence of the charges on the two colloids and maximum coagulation. However, Lottermoser in 1910 pointed out that it was better to consider the equivalence of the two stabilizing agents present. This was demonstrated by Freundlich and Nathanson in 1920, who showed that two colloids with *like* charges may mutually coagulate. Arsenic trisulfide is coagulated by Oden's sulfur, both being

negatively charged. The reaction between the respective stabilizing agents may be written as follows:

$$5H_2S + H_2S_5O_6 \rightarrow 610S + 6H_2O$$
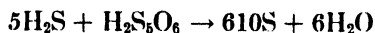
Thomas and Johnson in 1923 showed that with arsenic trisulfide and ferric hydroxide some reduction of the stabilizing agent of the latter occurs, according to the equation:

$$2FeCl_3 + H_2S \rightarrow 2FeCl_2 + S + 2HCl$$

They were able to show that the coagulum contained free sulfur. They also proved the equivalence of chloride ion and sodium ion involved in the mutual coagulation of ferric hydroxide and colloidal silica.

Mixtures of certain proteins taken in the proportions to produce an isoelectric mixture usually give a coagulum, but there are exceptions to this occurring at the isoelectric point. Hence Billitzer's suggestion is a factor, but Lottermoser's consideration of chemical equivalence of the stabilizing agents is overriding. Other examples of mutual coagulation are gums with gelatin chloride, or the coagulation of tannin with gelatin, where excess of gelatin is avoided. Another is the precipitation of silver sol by gelatin in just the right amount.

It should be noted that the above discussion applies equally to stable and unstable, lyophilic and lyophobic, reversible and irreversible colloids. For example, albumin can substitute for ferric hydroxide or chromic oxide in the coagulation of sols of silver, arsenic trisulfide, or silica.

The coagula, of indefinite composition, resulting from the mutual coagulation of two sols have often been taken for chemical compounds. The classical example is the Purple of Cassius. This was studied and then synthesized by Zsigmondy upon a mixture of sols of gold and stannic acid. Addition of the dye fuchsin to a sol of gold gives a coagulum not decomposed by water, but the dye may be extracted by treatment with alcohol leaving a black residue of gold, which upon pressing appears metallic. It is clear that there are innumerable examples of such "sorption compounds" and that chemical formulas are not properly applied to them, as in the case of the coagulum from ferric hydroxide and arsenic trioxide or of the result of mixing a basic and an acid dye. In general, the coagula are insoluble regardless of the solubility of one of the colloids alone. Such examples are mixtures of collodion with gelatin, casein, albumin, or oxyhemoglobin. The phenomena of coacervation will be dealt with later.

There are many applications of the mutual coagulation of colloids of opposite sign, probably the most important being the treatment of sewage and the clarification of turbid liquids.

Sensitization by Added Stable Colloids

If a small amount of albumin is added to a sol of ferric hydroxide in amount insufficient to produce coagulation, it is found that the ferric hydroxide sol is still positively charged. But it may be more sensitive to the coagulating influence of electrolytes as compared with the original hydroxide. A further phenomenon is that more concentrated sodium chloride may peptize the coagulum on account of the influence of the albumin. A similar behavior is exhibited by euglobulin added to dyes, including peptization by addition of enough sodium chloride. There are many examples of sensitization, such as the effect of gelatin or casein, of benzopurpurin, of tannin on many dyes irrespective of their charge, or of albumin upon gum mastic. But there are other instances where it does not occur, as for example with silver sol treated with gum arabic, saponin, tannin or dyes with albumin, or alkaline gold sols with gelatin.

It is clear that one must examine the relation of the two stabilizing agents involved, and that Billitzer's idea of mere electrical equivalence and Lottermoser's of chemical equivalence of the original stabilizing agent are not enough. The stabilizing agent itself undergoes an alteration. Kruyt [9] has shown that sensitization by tannin has nothing to do with the electrical charge of either the original colloid or the sensitized colloid, nor does it depend upon pH. Here a polar film of tannin surrounds the treated particles with the hydrophobic digalloyl groups directed outwards. Silver *protected* by excessive gelatin still retains a negative charge. Therefore the polar film exposes its negative ends. Sheppard [10] has emphasized this effect of orientation.

Silver halides can sorb dye molecules, the first layer of which have the hydrophilic ends adhering to the silver halide surface with the hydrophobic ends exposed to the aqueous solution. But further addition of dye, on the principle of like to like, gives another layer with the hydrophobic groups on the hydrophobic groups of the first layer, and the hydrophilic groups attached to the water. Similarly with silver halides treated with proteins. The new stability imparted by coating with excess of the stable colloid is thus explained and will be further discussed in the next paragraphs dealing with protective action. It is in these preliminary stages that sensitization is prominent.

Not only colloids but also nonelectrolytes, such as alcohol, may sensitize colloids, as was explained in the introduction to this chapter, by lowering the solubility of the stabilizing agent.

There are many important applications in medicine and pathology of the phenomena of sensitization. They are used for distinguishing hydrophilic colloids, and reference should be made to the Wasserman and Sachs-Georgi reactions, to the highly specific agglutinations of bacteria, and to the process of deproteinization by adding just the right amount of ferric hydroxide. The so-called iron number [11] measures over very wide limits the increased sensitization brought about by pathological and normal serums and the various proteins therein.

Protective Action of Stable Colloids

When a stable colloid is added in excess, not only is coagulation avoided but the new sol takes on many of the properties and much of the stability of the stable colloid. Faraday found that sufficient addition of gelatin to his colloidal gold enabled the latter to be evaporated and then redissolved in water. Bechhold [12] in 1904 is responsible for the explanation that the protective colloid operates by enveloping the original colloid with a film of the stable colloid and exposing its stabilizing agent. This process takes a few minutes to occur, and the concept explains the effect of the order of mixing the colloids, which might otherwise be perplexing. For example, acidified gelatin, gelatin chloride, added to ferric hydroxide and then treated with ammonium hydroxide would not give a coagulum. But when the gelatin was first treated with ammonium hydroxide and then added to the ferric hydroxide, coagulation was produced.

It is necessary for the stable colloid to remain stable in order to give protective action. For this reason albumin, casein, and edestin are poor protective colloids for cellulose because the product is too insoluble.

Striking demonstrations of protective action may be made by showing how a soluble soap can carry charcoal and finely divided ferric oxide through filter paper, when in the absence of the soap none will pass through. [13]

Another example is the precipitation of mercuric iodide in the presence of gelatin, which preserves the unstable yellow form, whereas with water the red form is almost instantaneously produced. [14]

There are more examples of practical applications of protective action than of any other property of colloids. After all, one of the chief

uses of soaps, detergents, and emulsifiers is to exert protective action on droplets and suspensions. The so-called emulsion of the photographic plate is silver bromide protected by gelatin. Many medicinal preparations are protected colloids.

A poor method of measuring protective action is the well-known *gold number*. This is determined by the amount of protective colloid that suffices to protect red gold under a given set of conditions (conc. NaCl). If no protective colloid is present the gold turns blue. If sufficient, the gold stays red. It takes very small amounts of gelatin to prevent the color change in the gold. The shortcoming of this method is that it gives only one number for one protective colloid regardless of its concentration and other conditions of use. For example, Spring [15] found that a 2% soap solution did not suspend finely divided charcoal any better than water, although a 1% solution was highly effective. Spring found the same kind of effects with soaps in other solvents such as alcohol. It may be noted that zinc or magnesium soaps or resinates are commonly employed as protective colloids in benzene or other hydrocarbon systems.

Another measure sometimes referred to is the *Congo Rubin number*. This is the amount of protective colloid in 100 gm of 0.01% Congo Rubin solution which almost suppresses the color change brought about by addition of 0.160 gm equivalents of potassium chloride in 10 minutes. This suffers from the same defect as the gold number method. Both also assume that what is a protective colloid for gold or Congo Rubin must be a protective colloid for everything else.

Protective colloids frequently prevent the appearance of precipitates, and they are often used to modify or prevent crystallization as in the manufacture of explosive lead azide.

COACERVATION

Some stable colloidal sols separate into two liquid layers, each containing some of the colloid when to the sol is added either another colloid or an electrolyte or a nonelectrolyte. Kruyt and Bungenberg de Jong have coined the term "coacervation" and include all these cases as examples thereof.[16] Some of the instances of mutual coagulation discussed above are claimed as coacervation. Indeed it sometimes happens that what is really the separation of a liquid layer looks at first sight like coagulation. "The lower layer may be considered as an accumulation of micellae only separated by a binary liquid. This explains the name 'coacervate' for the sublayers (acervus = heap)." [17] In some cases true equilibria are said to be involved.

The best general account of this subject has been given by Koets,[18] with well over 100 references. One group of cases is described by him in the following words.

When two hydrophilic colloids of opposite charge are mixed in solution, separation of a complex may be observed over a range of concentrations. In most cases this complex separates from the residual solvent (liquids in equilibria) in the form of small liquid drops, which unite to form a liquid, highly viscous layer (coacervation). In those complexes the attraction due to the opposition in charge is balanced by resistance due to the hydration of the two colloids. If the electric charge of one of the two components is too small no formation of a complex will take place.

Other treatises on the same subject are by H. G. Bungenberg de Jong.[19]

As a particular instance there may be cited the effect of adding gelatin, positively charged at pH 3, to gum arabic also at pH 3 and bearing a negative charge. If the turbidity is plotted against the proportions of the two hydrophilic colloids, the form of the curve is exactly that shown in the previous diagram for a mixture of oppositely charged colloids. (See Fig. 12·1.) In the middle there is a zone of maximum turbidity which practically disappears as either separate component is approached. At this point both density and viscosity are at a minimum, and they show but little electrokinetic movement. At the point of maximum turbidity the minute droplets which appear can form a liquid layer, the coacervate. The coacervate is shown to retain the positive and negative particles whose mutual attraction has brought them together by placing them in an electrical field and demonstrating that they can be electrolyzed apart. Some coacervates can be separated by diminishing the electrical field between them and by dispersing them again with acid, alkali, or mutual electrolytes, but they often become irreversible upon standing. The properties of the coacervate are considered to be due to a balance between the effect of the ionic layers and the hydration of the particles which are considered to be at a minimum in the coacervate.

In addition to such phase separation in dilute solutions of hydrophilic colloids containing particles of opposite electrical charge, the Dutch school recognizes coacervates derived from highly concentrated solutions of hydrophilic colloids of like charge and the tactoids and Schiller layers obtained with charged particles of anisotropic dimensions. In the first and third cases the particles are estimated to be at distances varying from 100 to 1000 Å from each other. The tactoids are swarms of elongated particles such as vanadium pentoxide, tobacco mosaic virus, myosin, fibrin, and benzopurpurin. Schiller layers form from disc-shaped particles such as the secondary aggregates of particles

of ferric hydroxide. Their streaming double refraction was referred to in Chapter 6.

The term coacervate has been broadened to include the interaction of a colloid and of an ordinary ion of opposite charge, usually polyvalent; for example, the coagulate formed from gum arabic or lecithin by addition of a medium amount of barium chloride, an excess of which redisperses the coagulate. Coagulates of this group are called "autocomplexes." The suggestion has been made that the protoplasmic membrane of the animal cell is a coacervate film. These conceptions have been stretched to include cases where no colloids are present, but only ordinary polyvalent ions.

REFERENCES

Text References

1. SEABRA, P., *Medical Times*, October, 1941
2. LUMIÈRE, A., *La Vie, la maladie, et la mort*, Masson, Paris (1928); DUBOIS, E. F., "The Mechanism of Heat Loss and Temperature Regulation," *Lane Medical Lectures*, Stanford University Press (1937)
3. FREUNDLICH, H., *Z. physik. Chem.*, 44, 136 (1903); KRUYT, H. R. and VAN DUIN, C. F., *Kolloidchem. Beihefte*, 5, 270 (1914)
4. WEISER, H. B. and MILLIGAN, W. O., *J. Am. Chem. Soc.*, 62, 1924 (1940)
5. BOLAM, T. R. and MUIR, J. J., *J. Chem. Soc.*, 765 (1934); BOLAM, T. R. and CURRIE, A. B., *ibid.*, 296 (1939); BOLAM, T. R. and DUNCAN, A. I. S., *ibid.*, 1317 (1936)
6. ANNETTS, M. and NEWMAN, L., *J. Phys. Chem.*, 40, 187 (1936)
7. USHER, F. L., *Trans. Farad. Soc.*, 21, 406 (1925)
8. RIDEAL, E. K., *Nature*, 152, 618 (1943)
9. KRUYT, H. R., *Kolloid-Z.*, 31, 338 (1922)
10. SHEPPARD, S. E., *Nature*, 155, 266 (1945)
11. WINDISCH, W. and BERMANN, V., *Wochenschr. f. Brauerei*, 37, 20 (1920)
12. BECHHOLD, H., *Z. physik. Chem.*, 48, 385 (1904)
13. SPRING, W., *Kolloid-Z.*, 4, 166 (1909)
14. FRIEND, J. N., *Nature*, 109, 341 (1922)
15. SPRING, *op. cit.*
16. KRUYT, H. R. and BUNGENBERG DE JONG, H. G., *Proc. Roy. Acad. Amsterdam*, 32, 849 (1929); *Kolloid-Z.*, 50, 39 (1930)
17. MODDERMAN, R. S. T. and HOLLEMAN, L. W. J., *Nature*, 129, 654 (1932)
18. KOETS, P. K., *Reports on Progress in Physics*, 10, 129 (1946)
19. BUNGENBERG DE JONG, H. G., *Généralités et coacervats complexes*, Actualités Scientifiques et Industrielles #397, Hermann et Cie, Paris (1936); *Coacervats autocomplexes*, #398, *ibid.*

ELECTRICAL AND ELECTROKINETIC

PHENOMENA

Bubbles, drops, and solids suspended in a liquid, including colloidal particles in sols and gels, exhibit bodily movement when placed in an electrical field. Therefore, they are electrically charged. Since the system as a whole is electrically neutral, there are an equal number of free positive and negative charges present. For each charge upon the surface or particle there is one of opposite sign on a free ion in the liquid. Therefore, since the movement of electrical charges of opposite sign past each other is synonymous with the passage of a current, all these systems must exhibit a small, but definite conductivity. Since these phenomena do not occur in nonionizing solvents, such as chloroform, dry ether, dry benzene, petroleum, but do occur distinctly in amyl alcohol, acetic acid, turpentine, and carbon disulfide, and freely in ionizing solvents such as methyl and ethyl alcohols, liquid ammonia, acetyl acetate, nitrobenzene, and water, it is evident that the phenomena have to do with ions.

Coehn [1] has pointed out that when two phases are in contact, the one with a higher dielectric constant tends to be positively charged, the other negatively charged. This rule holds, strictly speaking, for nonconductors, but it is of such wide applicability that since the dielectric constant of water is exceptionally high, namely 80, almost all colloidal particles, bubbles, and suspensions are negatively charged.

Conductivity

The conductivity of most colloidal solutions is very low, resembling that of distilled water. Careful measurements in rather pure solvent reveal the slight conductivity of the colloidal particles, although in some instances this is counterbalanced by their having sorbed some

of the impurities in the original solvent. Some other colloids, such as a ferric hydroxide sol that is stabilized by the partially hydrolyzed chloride, exhibit a correspondingly more pronounced conductivity. Finally, the whole group known as colloidal electrolytes (Chap. 17 and 18) exhibits a conductivity that is comparable with that of many electrolytes.

The most usual method of measuring surface conductivity is first to measure the conductivity of a fairly concentrated solution of electrolyte such as potassium chloride with a given surface or porous body such as a sintered membrane or mass of cellulose, assuming that the surface conductivity will be a negligible fraction thereof, and then to substitute a very dilute solution and ascribe to surface conductivity any conductivity above that calculated from the bulk conductivities of the two solutions.

Electrokinetic Phenomena

There are four chief ways in which the relation between motion and electrical charge or potential is exhibited, two of which depend upon the application of an electrical field and two upon the application of mechanical force. They are listed and defined as follows:

A. *External electromotive force.* (1) Electroosmosis (Reuss, 1808). This occurs when the solid is held so that it is not free to move, but then the liquid does. It is observed with immersed porous materials or compacted powders, with gels and jellies, and even in a heap of froth or between the two surfaces of a soap bubble or through human skin. In the latter case, the water carries with it any nonelectrolyte present, such as sugar or anesthetic.[2]

(2) Electrophoresis or cataphoresis. Here particles, animate or inanimate, including colloidal particles, drops and bubbles, move just like migrating ions.

B. *Mechanically caused motion.* (3) Streaming potential (Quincke, 1859, Zöllner, 1871). This results from forcing a liquid through any capillary system, such as a tube, membrane, porous plate, compacted powder, foam, gel, or jelly. It therefore occurs in ultrafiltration or in the seepage of water in soil. The potential often amounts to many volts.

(4) The Dorn (1878) effect. Potential arising from falling drops or particles or from rising bubbles, for example in an effervescent liquid.

Electroosmosis was systematically investigated by Wiedemann,[3] chiefly using a pot of porous earthenware. The work of Smoluchowski

(1903) and of Perrin (1904) should also be mentioned. The latter especially studied compacted powders. If the flow of the liquid is restricted, an "electroosmotic" pressure difference develops (R. Porret, 1816). The two chief rules deduced are: first, that for every voltage applied a proportional hydrostatic pressure is produced, and, second, that for every amount of current passed, a proportionate volume of liquid is transported.

The first instance may be visualized when a porous partition is placed in the bottom of an open U-tube partly full of liquid with an electrode inserted in each limb. Upon applying the e.m.f. the liquid rises on one side and falls on the other until it comes to a stationary position in which the difference in height is proportional to the voltage. The nature of this steady state is easily seen when the current is turned off, for the menisci immediately begin to equalize their levels. In the steady state the transportation of liquid towards the cathode is just balanced by the rate at which it leaks back, under the hydrostatic pressure, through the central portion of each pore or capillary.

The second instance is observed when the liquid is allowed to overflow freely, without opposing hydrostatic pressure, the volume of the liquid overflowing being proportional to the time as long as the current and hydrostatic pressure are left constant and the plug, in the electric field, acting as a steady pump. In other words, the flow is proportional to the total amount of current passed. The influence of electrolytes is discussed below.

√Electrophoresis (or cataphoresis) may be demonstrated with the same equipment as was used by Nernst to show the migration of the colored permanganate ion. The permanganate solution is placed at the bottom of a U-tube and is separated from the platinum electrodes that are placed near the top of the two limbs by a colorless solution of very dilute electrolyte. The band of color is then seen to move uniformly away from the cathode and toward the anode. Similarly, if colloidal ferric hydroxide is substituted for the permanganate, the brown color moves toward the cathode. The ferric hydroxide is positively charged, the corresponding negative charges being the chlorine ions. The migration in both cases may be measured quantitatively by analysis of the substance in the anode half or the cathode half, or both, before and after a known amount of current has passed. The sole proviso is that a middle portion of unaltered composition must be preserved, not allowing the processes on the two sides long enough time to extend their influences until they meet. No membrane should be used because of the electroosmotic complication. By taking suitable

precautions [4] one arrives at a necessarily identical result from merely measuring the linear rate of movement of the boundary and the slope of potential in which it is moving. The identity follows from the conservation of matter. Sometimes the movement is observed under the microscope or ultramicroscope and here again an identical result must be obtained if the proper precautions are observed.[5] Here the observation tube or cell should be of rectangular cross section, and, to eliminate the effect of electroosmosis in its steady state, the particles should be observed in the layer at 20% or 80% of the distance between opposite walls.

The linear rate of movement of proteins has been intensively studied and highly refined in connection with the characterization and investigation of proteins, as described in Chapter 18.

Electrodialysis. Qualitatively the direction of electrophoresis is conveniently observed by placing the liquid in a siphon the ends of which are closed by a membrane, such as parchment or Cellophane, dipping into a cathode and anode vessel respectively. The particles cannot move through the membrane and tend to be deposited on the one toward which they are moving. Meanwhile all ordinary anions and cations are migrating through toward the electrodes. This is termed electrodialysis and is by far the most thorough way of purifying a colloidal system from electrolytic impurities, although ion exchange is far quicker. In practice the central compartment is made compact, and the liquid around the electrodes is constantly replaced with pure, fresh solvent. Since the colloid is accumulating near one membrane, the process of purification, as well as of concentration, is greatly hastened by pouring off the rest of the liquid from time to time and replacing it by pure solvent. This is Pauli's electrodecantation.[6]

Sign of Charge and Influence of Electrolytes

Colloid particles which migrate toward the cathode and are positively charged include the well known ferric hydroxide sols and those of the oxides or hydroxides of aluminum, chromium, titanium, zinc, cobalt, and cadmium; the carbonates of barium, strontium, and sometimes calcium; the oxidized sols of the base metals prepared by Bredig's method, iron, copper, bismuth, and lead; the remarkable insoluble crystals of chromic chloride; solutions of the basic dyes and proteins in acid solution.

Most materials migrate toward the anode and are negatively charged. These include the common colloids, silicic acid, vanadium

pentoxide, starch, oil, the gums, iodoform in particular, charcoal, the acid dyes, and the proteins in alkaline solution. Under special conditions most colloids may be prepared with an unusual sign of charge, as Frumkin's positive platinum sols.

Hardy in 1900 emphasized the outstanding importance of hydrogen and hydroxyl ions in imparting their own charges (by being sorbed) to many colloids. The effect of progressive addition to a particle of opposite charge is to diminish its movement and charge, usually passing through a range of instability and often then reversing and building up the new sign of charge. Only two other ions need specific mention for special effectiveness, the silver ion and the sulfosalicylate ion; thus colloids treated with sulfosalicylic acid usually stay negative.

Returning for a moment to a reconsideration of the demonstration of electrophoresis, we now see why the colloid should be protected from the influence of the electrode solutions where alkali is being produced at the cathode and acid at the anode, each with a tendency to diminish or reverse the charge of the particles travelling to them and toward which they themselves migrate. The final state could thus be complete coagulation.

The effect of higher valency is outstanding, and it is the ion of opposite charge sign to that of the particle that is more effective. In general, electrolytes are better at discharging than at recharging. These phenomena are closely similar to those discussed in Chapter 12 on Coagulation, Table 12.

Rate of Migration

Hydrogen and hydroxyl ions are the fastest ions known. They move at the rate of $35 \times$ and 18×10^{-4} cm/sec, respectively, in a field of 1 volt/cm. Ordinary fast ions are those of potassium and chlorine whose movement is at the rate of 6.8×10^{-4}. The heaviest and slowest ions move at least 2×10^{-4} cm /sec /volt /cm.

Now it is an extraordinary fact that bubbles, drops, suspensions (animate and inanimate), and colloidal particles move at the same rate as ordinary ions, that is, between 2×10^{-4} and 20×10^{-4} cm /sec / volt /cm. Thus a negative charge riding on a drop of paraffin oil may move much faster than if it were being carried by an atom of chlorine. In other words, this electron on the paraffin surface is moving or conducting better than a chloride ion. However, the total conductivity of the colloidal solution is usually low because the total number of particles and charges is small as compared with the total number of

ions in an ordinary solution of an electrolyte. The number of charges is usually only a minute fraction of the number of chemical equivalents of the charged body.

The magnitude of the charge has been a matter of curious uncertainty, various modes of calculating it for the same system having differed by as much as 100,000-fold.[7] It ranges from such a low value as one charge per 5000 nickel atoms in Hatschek's nickel sol, stabilized by rubber in benzene, to one charge per chemical equivalent in the true ionic micelle of the author's soap solutions and sulfonic acids. Here again rate of movement and actual conductivity is equal to that of the same charges on a fast ion. The charge on a ferric hydroxide sol has been determined [8] as 1 positive charge to 206 chemical equivalents of ferric iron, the rate of movement and conductivity being 57% of that of the same charge on potassium ions. The corresponding free chloride ion therefore carries $1/1.57$ or 64% of the current, leaving 36% for the colloidal particles.

Origin and Nature of Charges

It has been emphasized, especially by Freundlich, that the electrokinetic charge on a surface or a colloidal particle bears little or no relation to the ordinary electrode charges or Nernst potentials. For example, a mercury electrode is usually positively charged, but an emulsion or sol of mercury is negative. The charges may be derived from the dissociation of surface molecules as when glass in contact with water dissociates off a sodium or calcium ion leaving a negative charge upon the glass. This is also the source of charges on a soap particle. Alternatively, the charges may arise from dissociation of molecules adsorbed on the particle. This was illustrated in Chapter 1, Figure 1·2, in the partial dissociation of hydrogen ions from hexahydroxyplatinic acid coating and stabilizing the surface of a platinum particle. A third alternative is the preferential sorption of one ion rather than another, as for example on a paraffin surface. A fourth alternative which shows how complicated results of these simple effects may be is the case pointed out by Mukherjee of a negatively charged particle which sorbs trivalent aluminum ion, this excess of positive charges then sorbing a chloride ion. The total effective charge on any particle or a surface is the net algebraic sum of all the charges thereon. If the opposite charges are equal, as with a protein or clay at its isoelectric point, the behavior is that of a zwitter ion. Another kind of isoelectric surface would be one having no charge at all.

The Relation of Electrokinetics to Electrochemistry

Apart from a few exceptions, the conduction of electricity is electrolytic and lies in the field of electrochemistry. The usual exceptions are the cases of metallic or electronic conduction, as in the metals and in a few similar crystalline materials, such as magnetite. Two other exceptions are also to be noted, one being the solutions of metals in liquid ammonia where conduction is considered by Kraus to be partly electrolytic and partly metallic or electronic, as by the movement of solvated electrons toward the anode. Solvated electrons are also considered by many physicists to be at least partly responsible for the phenomena of waterfall electricity and the charges on the air-water interface.

Otherwise the conduction of electricity is electrolytic. It consists of the movement of charged matter past other charged matter of opposite sign, and to this process the laws of Ohm and Faraday apply. The simplest cases are the simple ions like those of potassium and chlorine where there is one charge on one atom and the charge is invariable. When the charge varies with polyvalent atoms, such as calcium or cadmium, one postulates [9] a formation of so-called intermediate monovalent ions such as CdI^+ and $C_6H_5 \cdot CH(OH) \cdot CO_2Ca^+$. Other complex ions are common, an example [10] being CdI_3^-. A surface or a particle may carry many charges at various points, not all of the same sign and not invariant. All electrokinetic movement, however, produces its corresponding amount of electrolysis.

Electrical forces are enormous as compared with ordinary mechanical effects. For example, one equivalent or Faraday of electronic charges in a field of only 1 volt/cm experiences a force upon it equal to the product of the charge and the potential gradient, amounting to 1 million kilograms, being $4.774 \times 10^{-10} \times 6.06 \times \dfrac{10^{23}}{300}$ dynes. The force required to move 1 gm of hydrogen ion with a velocity of 1 cm/sec is 300 million kilograms. Yet a solution of potassium chloride, or of hydrochloric acid, or of any colloid, lies apparently quiescent in the bottom of an open U-tube when a current is being passed through it. This is because the total friction on the positive charges is precisely equal to that on the negative charges. The current is distributed amongst all the charges exactly in proportion to the conductivity of each, including even the charges on the surface of the glass tube. Each moves past the liquid, or the neighboring liquid moves past it, with a velocity which is inversely proportional to its friction or directly

proportional to its mobility. Measurement of the distribution of the current among the available carriers is an excellent method of determining their respective conductivities [11] and of comparing them with those of known ions. It has been shown [12] that the general formula

$$n_1 = c_1 m_1 f_1 / \mu$$

expresses the actual movement, relative to the solvent, of each constituent of any electrolytical conducting system, where n_1 is the migration or transport number of constituent 1, c_1 its concentration in chemical equivalents per 1000 gm of solvent, f_1 its conductivity contributed per chemical equivalent, m_1 the number of chemical equivalents of constituent 1 per Faraday of free charge, and μ the total conductivity of that amount of solution that contains 1000 gm of solvent, and equals the sum of all the terms $c \cdot f$; that is, the concentration of each constituent multiplied by its mobility. Conversely, the movement of the solvent relative to any conducting constituent whatsoever (electroosmosis) is $m_1 f_1 / \mu$ kilos of solvent per Faraday of current. It was from this equation that the value m_{F_e} was found to be 206 equivalents of iron for each positive charge on the "ferric hydroxide" particle in the example quoted in a previous paragraph.

The only distinction between electrokinetic and ordinary ionic movement appears to be that in some of the phenomena of the former, namely, electroosmosis and streaming potential, we have the ability to hold in place some of the charged matter. It will be seen that the mechanical forces involved are insignificant in comparison with the electrical stress exerted equally upon the opposite charges. All the phenomena of electrokinetics, if referred to movement past the solvent are identical with those common to the whole of electrochemistry.

"Electrokinetic or Zeta Potential"

Many authorities do not consider that the approach to this subject from general electrochemistry should prove especially fruitful, whether valid or not. The author considers that electrochemistry embraces the whole subject, both qualitatively and quantitatively. Others prefer to build upon the mathematical basis provided by Helmholtz in 1879, which antedated any knowledge of the dissociation of electrolytes. The subject has been modernized by combining the double layer of Helmholtz, which is regarded as an electrical condenser, with the diffuse ionic layer, which was first developed by Gouy in 1910 and is now familiar in the ionic atmosphere of the Debye-Hückel theory of

electrolytes and in the Donnan equilibrium. The electrokinetic phenomena are now usually ascribed to this diffuse outer layer or a portion thereof, and there would be nothing to criticize in the calculations [13] provided that their basis corresponds to the actual facts of the case.

The author's belief that the results are illusory or erroneous rests upon two grounds. The chief reason is that the charges upon surfaces and particles are spaced so far apart that they are practically independent of each other. For example, in a decinormal solution of a simple electrolyte the ions are at an average distance of only 20 Å apart and yet this suffices to leave them, as a first approximation, as separate individual ions. Similar distances prevail between the charges even on the highly charged ionic micelle of soaps. A close-packed monomolecular layer of ions could lead to potential differences across the interface of 56 volts, and Helmholtz, Lamb, Cameron and Ottinger considered that the actual voltage between a glass tube or a porous pot and an aqueous solution in contact with it was as much as 5 or more volts. Such voltages are definitely no longer under consideration and electrokinetic potentials are now usually reckoned in millivolts. The mathematical treatment according to which the electrokinetic potential ζ is derived assumes a continuous uniform electrification, a continuous smear, whereas the considerations just adduced lead to a distribution so discontinuous that each charge would possess its independent ionic atmosphere just as does each charge in a decinormal solution of potassium chloride. Frequently also, the charges on a given surface or particle are not all of the same sign.

The electrokinetic potential ζ has never been measured but always calculated from observed movement or from a formula derived from observed movements. By definition the ζ potential in millivolts is the linear movement in cm/sec/volt/cm multiplied by $4\pi\eta/D$, where η is viscosity and D is the dielectric constant. This factor (Helmholtz-Smoluchowski) equals 148,400 at 18°, 142,600 at 20° and 129,700 at 25°. Some writers have used instead the Debye-Hückel factor $6\pi\eta/D$. The same operation can be applied to the mobility of any ion, resulting in the value -102.6 millivolts for an ordinary chloride ion at infinite dilution as compared with $+58.6$ millivolts for the corresponding positive particles in a "ferric hydroxide" sol.[14]

A second less fundamental reason for the author's disbelief in the reality of the electrokinetic potential either for ions or for particles except merely as an alternative numerical method of expressing actual mobilities or linear velocities of electroosmosis (the two are identical) is that the conductivity observed for surfaces is too great to be ac-

counted for by an outer diffuse ionic atmosphere.[15] For the most recent formulation of the electrokinetic effects the reader is referred to the publications of Booth [16] and Elton.[17]

Besides the usual calculation of ζ potential from electroosmosis or cataphoresis,

$$\zeta = 4\pi\eta U/ED$$

where E is the electric field and U is the linear velocity of liquid past solid or solid past liquid, we have

$$\zeta = \pi r^2 P/2ED$$

for the electroosmotic pressure P in the steady state, and

$$\zeta = 4\pi\eta E\lambda/DP$$

for the streaming potential E where P is the applied hydrostatic pressure and λ is the specific conductivity of the liquid.

Discussions of the conductivity produced by the electric double layer on a charged surface now discriminate between, on the one hand, sessile charges which are adsorbed on the surface and are directly in contact with it and which form the outer layer of Helmholtz double layer, and on the other hand, the mobile ions which are in the diffuse ionic atmosphere originally postulated by Gouy upon considerations of thermal vibration. The mobile ions constitute a small fraction of the "fixed, bound, or sessile" ions.[18]

It is clear that if a double layer on a single surface can thus release enough of the ions of opposite charge to the surface (gegenions) to produce a diffuse ionic layer extending out to a depth of hundreds or thousands of Ångström units, when two such surfaces of like charge are opposed to each other with an equal number of ions of opposite charge disposed on or between them so as to make the whole system electrically neutral, these two surfaces will be thrust a corresponding distance apart. This pushing apart of like charged surfaces in spite of the gegenions between them has been the subject of calculation by Langmuir, Derjaguin, Verwey, Levine, and Schofield.[19] They appeared to differ but Brady [20] has recently recalculated them and shown that their results are mathematically interconvertible. Brady has further introduced the assumption that the surfaces will be pushed at least so far apart that the osmotic coefficient or ionic strength of the diffuse double layer half way between them has fallen to the value obtaining throughout the whole liquid.

Evidently this is a problem which is fundamental to discussing the stability of charged colloids.

This treatment is incomplete because it supplies only the reasons why the surfaces are pushed a definite distance apart without explaining why they do not then become completely independent of each other and drift farther away. Now there are colloidal structures such as Schiller layers and the lamellar micelles discussed in Chapter 17 where the surfaces are spaced a definite distance apart, no more and no less, although the exact value may change with concentration and addition of electrolytes. This necessitates the existence of some kind of balancing attractive forces which Hamaker [21] and the Dutch school have sought in very long range attraction by van der Waals' and London's forces. Even this, however, cannot complete the whole story because even when addition of an electrolyte has the predicted effect in very dilute solution, the effect may be in the opposite direction for less dilute solutions.

There is therefore more to be discovered in this subject. It may be mentioned that Holliday [22] finds that during the dialysis of gold sols which turn from red to blue, the particle charge rather than the ζ potential is the critical factor. It may be added that surface conductivity and ζ potential are not in general closely related.[23]

Electrocapillarity

For a discussion of electrocapillarity the reader must be referred to one of the recent reviews of the subject such as that of Graham [24] or to one of several standard treatises. However, it should be mentioned that Frumkin [25] has described an electrocapillary movement of liquid mercury droplets that is several orders of magnitude faster than electrokinetic movement.

Applications

There are many applications of electrokinetic processes and especially of electroosmosis, besides the electrodialysis already referred to for the purification and the characterization of sols. Electroosmosis has been used for the purification of china clay, reducing the water content to as little as 18% and, under the best conditions, the current required to deposit 1 gm of hydrogen will deposit 700,000 gm of clay. Peat is dewatered similarly. Pure silicic acids are purified for therapeutic purposes by such processes; likewise antidiphtheria serums and antitoxins. Electrodeposition of rubber latex on forms is another industrial application. The localization of surface active agents in electro-

plating, the separation of oil from water, as in turbine condensate emulsions, and the concentration of ores in clay deposits, and many other processes, are all examples of the application of electrokinetics in industry and science. A number of clinical applications have been summarized by Abramson and Gorin.[26]

REFERENCES

Electrical and Electrokinetic Phenomena

ABRAMSON, H. A., *Electrokinetic Phenomena and Their Application to Biology and Medicine*, A.C.S. Monograph No. 66, Chemical Catalog Co., New York (1934)

MCBAIN, J. W., *J. Phys. Chem.*, 28, 706 (1924)

Hydrophobic Colloids, A Symposium, D. B. CENTENS UITG MIJ, N. V. Amsterdam (1937)

VERWEY, E. J. W. and OVERBEEK, J. T. G., *Theory of the Stability of Lyophobic Colloids*, Elsevier Publishing Co., Inc., New York (1948)

Text References

1. COEHN, A., *Ergebnisse der exakten Naturwissen.*, 1, 175 (1922)
2. ABRAMSON, H. A. and MOYER, L. S., *Trans. Electrochem. Soc.*, 1937, 135
3. WIEDEMANN, G., *Pogg. Ann.*, 87, 321 (1852)
4. MCBAIN, J. W. and THOMAS, W. M., *J. Phys. Chem.*, 40, 997 (1936), and the references therein
5. ABRAMSON and MOYER, *loc. cit.*
   MCBAIN and THOMAS, *loc. cit.*
6. STAMBERGER, PAUL, *J. Coll. Sci.*, 1, 93 (1946)
7. USHER, F. L., *Trans. Farad. Soc.*, 21, 415 (1925)
8. MCBAIN and THOMAS, *loc. cit.*
9. DAVIES, C. W., *J. Chem. Soc.*, 271 (1938)
10. MCBAIN, J. W., *Z. Elektrochem.*, 11, 215 (1905)
11. MCBAIN, M. E. L., *Trans. Farad. Soc.*, 31, 153 (1935)
12. LAING, M. E., *J. Phys. Chem.*, 28, 673 (1924)
13. For example, MULLER, H., *Cold Springs Harbor Symposia on Quantitative Biology* (1933) p. 9; BRIGGS, D. R., *ibid.*, p. 14
14. MCBAIN and THOMAS, *loc. cit.*
15. MCBAIN, J. W., PEAKER, C. R. and KING, A. M., *J. Am. Chem. Soc.*, 51, 3294 (1929); MCBAIN, J. W. and PEAKER, C. R., *Proc. Roy. Soc.*, 125A, 394 (1929); MCBAIN, J. W. and FOSTER, J. F., *Colloid Symp.* (1934); *J. Phys. Chem.*, 39, 331 (1935); RUTGERS, A. J., VERLENDE, ED. and MOORKENS, MA., *Proc. Koninklijke Nederland Akad. van Wetenschappen*, 41, 163 (1938); RUTGERS, A. J., *Trans. Farad. Soc.*, 36, 69 (1940)
16. BOOTH, F., *Nature*, 161, 83 (1948)
17. ELTON, G. A. H., *Nature*, 161, 847 (1948); *Proc. Roy. Soc.*, 194, 259 (1948)
18. STERN, O., *Z. Elektrochem.*, 40, 508 (1924)
19. LANGMUIR, I., *J. Chem. Phys.*, 6, 873 (1938)
    DERJAGUIN, B. and LANDAU, L., *Acta Physicochim.*, 14, 633 (1941)
    VERWEY, E. J. W. and OVERBEEK, J. T. G., *Trans. Farad. Soc.*, 42B, 117 (1946)
    LEVINE, S. and DUBE, G. P., *Trans. Farad. Soc.*, 35, 1125, 1141 (1939); 36, 215 (1940)
    SCHOFIELD, R. K., *Trans. Farad. Soc.*, 42B, 219 (1946)

20. BRADY, A. P., *22nd Annual Colloid Symp.*, M.I.T. (June 1948)
21. HAMAKER, H., *Physica*, 10, 1058 (1937)
22. HOLLIDAY, A. K., *Trans. Farad. Soc.*, 43, 661 (1947)
23. RABINOV, G. and HEYMANN, E., *J. Phys. Chem.*, 47, 655 (1943); FAIRBROTHER, F. and BALKIN, M., *J. Chem. Soc.*, 1564 (1931)
24. GRAHAM, D. C., *Chem. Rev.*, 41, 441–501 (1947)
25. FRUMKIN, A., *J. Coll. Sci.*, 1, 277 (1946)
26. ABRAMSON, H. A. and GORIN, M. H., *Chem. Prods.* (London), 3(4), 37 (1940, February)

DIFFUSION

Diffusion is one of the simplest and yet most significant properties of substances in solution. For example, with electrolytes the two chief effects held responsible for changing the conductivity from the intrinsic value observed at infinite dilution, namely, "relaxation of the ionic atmosphere" and "electrophoretic" or viscous drag, are almost eliminated because in diffusion the two ions are moving in the same direction. Hence, any deviations occurring in real solutions as compared with infinite dilution have direct significance. Until recently, little attention could be given to the results of experiments on diffusion because the experimental difficulties were such that a very arbitrary choice was required in any attempt to deduce quantitative information from the conflicting experimental data obtained at great expense of time and trouble.

The diffusion of colloids differs only in degree from that of molecules and electrolytes and the rates of diffusion are often of similar magnitude.

The usual method has been some variant of the following procedure. The solution is placed in the bottom of a vessel and upon it is laid, without mixing, three volumes of solvent. After a lapse of a definite time (days or weeks) the four consecutive volumes are drawn off one by one and analyzed.

Stefan in 1879 drew up tables [1] of distribution amongst the four layers, by comparison with which a single result is yielded. The analysis may be made by an optical method without interrupting the diffusion.[2] Cohen has repeatedly emphasized the almost unavoidable disturbances due to vibration and to temperature fluctuations. In any case the results for potassium chloride, the best studied substance, were highly contradictory.[3]

The device introduced by Northrop and Anson [4] of arranging that diffusion should take place between two homogeneous bodies of liquid through a porous partition of sintered glass (or of alundum) provides a very simple, rapid, and reproducible method of obtaining diffusion data for electrolytes, nonelectrolytes, mixtures, and colloids under very varied conditions.
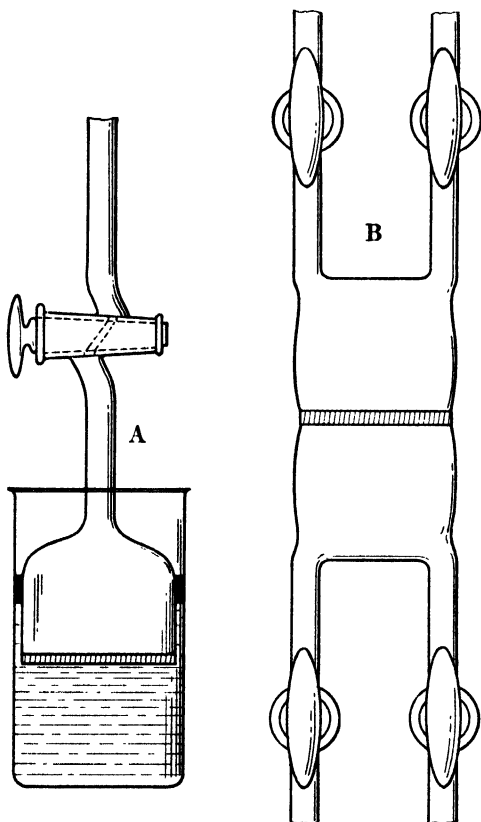
FIGURE 14·1 Porous disc diffusion cells. A, simplest form; B, for diffusion of any solution against any other solution.

With the sintered membrane method [5] (Figure 14·1), the heavy solution is placed above the membrane and the lighter one or pure solvent below. After a lapse of time the diffusion columns are set up. Then without disturbing them, both liquids are replaced and the experiment proper begins. Gravity usually suffices to keep both liquids homogeneous, although some authors [6] prefer to use direct stirring. The experiment is completed within twenty-four hours, but the results should be found independent of time and reproducible to a few tenths of one per cent. Since nothing is known of the number and shape of the pores containing the diffusion columns, the cell has to be standard-

ized, just as in conductivity determinations, by measuring a known standard [7] such as 0.1 N potassium chloride whose integral diffusion coefficient D at 25° C is 1.838×10^{-5} cm²/sec or 1.588 cm²/day. This yields a cell constant K whose significance is purely geometrical. The actual measurement is calculated by the formula:

$$KD = \frac{\log_{10} c_0 - \log_{10} (c_0 - 2c)}{t}$$

where D is the diffusion coefficient of the substance in question, c_0 is the initial and c the final concentration of the more dilute solution after the lapse of time t.

The method was formerly criticized or rejected by several authors on the following grounds: first, because the solution is inadequately stirred, and secondly, because it is supposed that the diffusion through the irregular pores might be complicated by slight convection within them due to a tendency for regions of equal density to lie horizontally across each pore.

In the first place, insofar as these effects have a geometrical basis, they have been eliminated by the method of obtaining the cell constant. In the second place, the membranes have to be tested for the complete absence of flow under a far higher hydrostatic pressure, namely, when the whole membrane is tilted by 20 or 30 degrees from the horizontal without affecting the observed rate of diffusion. Finally, the ratios between diffusion coefficients of potassium chloride and other substances, such as sucrose and hydrochloric acid, obtained with the sintered membrane method agree with the best existing data using the classical method. All results are based on the same standard solution and if that or any other substance is ever established more accurately, they may all be corrected by the corresponding factor.

Harned [8] has pointed out that the only confirmation of the Nernst equation has been with dilute aqueous solutions of potassium chloride using this sintered diaphragm cell method and that it is the only accurate method for dilute solutions and the only one that has confirmed the diffusion theory, but that for concentrated solutions an optical method such as that of Lamm and Polson [9] may be easier and preferable. Gordon has assembled the table on page 203, comparing the results obtained by McBain and Dawson, Gordon, and Hartley and Runnicles, showing a rather remarkable agreement.

The classical method has been refined by Tiselius and Gross [10] and by Lamm and Polson,[11] using the optical methods of observation developed in connection with use of the ultracentrifuge (Chapter 16).

TABLE 15

DIFFUSION OF AQUEOUS POTASSIUM CHLORIDE SOLUTIONS

$(C_o'' = 0; V' = VII; C_t/C_o = 0.5)$

C_o'	McBain & Dawson* D_i/D_o	Gordon D_i/D_o	Hartley & Runnicles** G-4 D_i/D_o	G-3 D_i/D_o
0.10 N	0.918	0.918	0.918	
0.05		0.935	0.934	
0.025		0.952	0.950	
0.02	0.949	0.956		
0.01	0.966	0.967	0.968	0.968
0.005			(0.980)	0.977
0.0025			(0.990)	0.984

* McBain and Dawson, *loc. cit.*
** Hartley and Runnicles, *Proc. Roy. Soc.*, **168A**, 401 (1938)

A very rapid optical method of determining diffusion by observing the blurring of a sharp boundary has been recently developed.[12]

The general equation for diffusion was adapted by Fick in 1855 from the equation for the conduction of heat. Fick's diffusion law may be expressed in either of two exactly equivalent forms:

The amount of substance ds diffusing through a unit area in time dt is proportional to the gradient of concentration c with linear distance x, thus

$$ds/dt = D \cdot dc/dx$$

The rate of concentration change at any given point is

$$dc/dt = D \cdot dc^2/dx^2$$

D is assumed to be a constant independent of concentration, an assumption that is, however, only a first approximation.

Deviations from constancy are of direct significance. McBain and Liu [13] point out how remarkable it is that the value for sucrose is 0.462 in 0.05 N and 0.378 in 1.0 N solution, a difference of only 18%, whereas the viscosity η of the two solutions is 1.047 and 3.080, respectively, or a difference of over 300%. Evidently when sucrose molecules are diffusing in the same direction, any retardation of one by colliding with another is made up by the acceleration of the other. The effective viscosity is that of the solvent.

Only for a diffusion of a third substance through a solution would the viscosity of the solution as distinguished from that of the solvent be a determining factor.

Sutherland [14] in 1904 and Einstein [15] in 1906 deduced for uncharged, independent spherical particles such as molecules of sucrose, large in comparison with those of water, and with activity coefficient equal to unity, the formula:

$$D = \frac{RT}{N} \cdot \frac{1}{f}$$

where f is the friction upon each particle, or since Stokes' law should hold for such uncharged spheres, $f = 6\pi\eta r$, or

$$D = \frac{RT}{N} \cdot \frac{1}{6\pi\eta r}$$

This may be used for determining size of monodisperse particles.

The only other authenticated formula is that of Nernst [16] for uni-univalent electrolytes:

$$D = \frac{2\,RT}{\dfrac{1}{u} + \dfrac{1}{v}}$$

Although this was deduced for infinite dilution, it still gives a surprisingly good approximation in real solutions if the van't Hoff iota value (this is twice the osmotic coefficient g) is substituted for the 2 in the numerator. For more general equations see Dean,[17] Gordon,[18] and Mrs. McBain.[19] Strictly speaking this equation then applies to the differential or so-called "instantaneous" diffusion coefficient at a precise concentration, whereas most measurements are made by diffusing a solution into pure water, giving an integral coefficient over all intermediate values down to zero. The relation between the two is:

$$D_{\text{diff}} = D_{\text{int}} + c \cdot dD_{\text{int}}/dc$$

where c is the concentration. Thus the integral and differential coefficients may be calculated from each other if the data are sufficiently numerous or accurate.

Self-diffusion [20] is a new method rendered possible by the use of radioactive isotopes, which permits the study of a single ion or other constituent and allows it to be followed down to such low concentrations as 10^{-5} N. The radioactive ion simply exchanges places with the ordinary ion moving in the opposite direction through an otherwise uniform concentration.

It is interesting to quote a few examples of diffusion coefficients which will disabuse the reader of the idea that colloids cannot diffuse. This common misconception is due to the early experiments, which are still valid, using a jelly or membrane to hold back colloidal

particles and prevent them from diffusing, letting only simple ions and molecules through. This was ultrafiltration, not free diffusion, but is of importance in life processes where sorption and other distortions also occur:

N/10 hydrochloric acid	2.71	cm²/day
N/10 potassium chloride	1.588	"
Sucrose	0.46	"
Isoelectric egg albumin	0.10	"
Egg albumin at pH 8	0.41	"
0.2 M potassium laurate	0.47	"
0.15 M sodium oleate	0.28	"
Ferric hydroxide at pH 5	0.61	"

Where the colloidal particles are charged, in pure water they and their gegenions must move at the same rate to preserve electroneutrality. The resulting rate of diffusion is an average between the inherent values which the colloidal particles and the ions would show separately if each were free to move independently. To minimize the electrical effects of the gegenions, an excess of electrolyte such as potassium chloride is added. However, putting in a buffer or a foreign electrolyte such as potassium chloride is always attended with the risk of altering or displacing the original stabilizing agent.

All charged particles and ions mutually affect each other's rate of diffusion. Vinograd and McBain [21] have given the integral form of the diffusion equation for any one positively charged ion or particle in a mixture of salts, all diffusing into pure water. For anions the minus sign must be changed to plus:

$$D_+ = u_+ RT/z \left[1 - z \frac{\Sigma u_+ c_+ /z_+ - \Sigma u_- c_- /z_-}{\Sigma u_\pm c_\pm} \right]$$

where z is the valence of the ion without regard for sign, u its equivalent mobility, c its equivalent concentration, R the gas constant = 6.66, D in cm²/sec and T the absolute temperature. Multiplying this value of D by 86,400 converts it to cm²/day.

We may conclude from this equation that ions move under two forces, their concentration gradient and the electrical gradient, because of the unequal diffusion of various charged particles and ions. A fast diffusing ion retards all other ions of the same sign, but accelerates all ions of the opposite sign. A cation present in relatively small amount will be accelerated by the accompanying diffusion of a salt that has a slower cation than anion and will be retarded by a salt whose cation is faster than its anion. For example, the fast hydrogen ion itself can be made by the presence of barium chloride to diffuse

far faster even than if it were free to diffuse at its own intrinsic value of 8.05 cm²/day, and a slow cation may be almost completely stopped by the electrical effect opposing the concentration gradient. If some electrolyte such as potassium chloride is initially present at uniform concentration throughout the system in which the diffusion takes place, these effects are progressively suppressed, depending almost entirely on the conductivity of the added uniform electrolyte. However, the concentration of this electrolyte ceases to be uniform — giving a "negative diffusion coefficient." To suppress the electrical effects by 95% the uniform electrolyte must have at least ten times the conductivity of the diffusing electrolyte.[22]

An extreme example of acceleration of diffusion is provided [23] by the mutual diffusion of a solution of glycine and a solution of hydrochloric acid into each other. Here the chloride ion can be made to move as fast as 7.3 cm²/day, which is five times faster than its normal velocity, and glycine can be accelerated to three times its normal rate. Hence in reactions in biological systems the rate of transfer of material may be many times as rapid as would be predicted from ordinary diffusion data.

Periodic Precipitation and Liesegang Rings

When a salt is diffusing through a medium containing another salt with which it may react to form a precipitate, it might have been assumed that the precipitate would be formed as a continuous process resulting in a fairly uniform, steadily increasing zone of precipitation. However, this is frequently not the case. The precipitate usually forms first in a narrow band and then later and further on in a series of such bands, which may be very numerous. These are usually known as Liesegang rings, although the phenomenon of periodic precipitation in tissues was known before his time.[24]

The usual procedure consists in allowing gelatin solution containing one of the reactants such as potassium dichromate to set to a gel in a test tube and then placing upon it a solution of the other reactants such as silver nitrate, whereupon a series of bands proceeds to form at increasing distances below the surface of the gelatin. Alternatively, a drop of the reactant can be placed upon a sheet of gelatin, whereupon a series of concentric rings begins to form.

However, the presence of colloid is not essential for the production of such periodic precipitation. It can occur wherever convection is avoided, as in capillaries, between sheets of glass, and in porous media.[25]

Morse prepared forty different precipitates with solutions between glass plates. Wilhelm Ostwald first accounted for rhythmic banding by assuming a supersaturation up to a certain point, followed by a release through rapid precipitation. Then, further on, supersaturation is built up again. Morse pointed out that slightly supersaturated solutions may persist for an indefinite time, that highly supersaturated solutions can last for only a short time before precipitation, and that this accounts for the rhythmic bands.

Although colloids and gelatins are not essential for producing these phenomena, the jelly can play a part through sorption and other interaction within the reactants or precipitate. Thus, Hatschek showed that periodic precipitation can result in jellies already containing crystals of precipitate which appear to be without influence. Presumably, such crystals are coated and protected by a film of gelatin.

The subject of periodic precipitation has a voluminous literature, including several books, because such phenomena are a common occurrence in plants and other organisms and in mineralogical and geological formations. Thus Hatschek was able to imitate the formation of all the naturally occurring forms of gold, and he likewise was able to get banded structures such as agates. A particularly beautiful example was obtained by Morse in which each ring consists of a single dendritic crystal of alum.

Diffusion through Sieve Membranes

Membranes set up barriers. If the pores are fine enough to allow only the solvent molecules to pass, they are "semipermeable." If the pores are larger they may hold back only colloidal particles, an effect used in dialysis. Even if the pores are rather coarse, if the membrane is charged (that is, if the membrane has fixed ions of one charge combined with mobile gegenions of the opposite charge), the effect of the charge on both ions will be felt by all ions and charged particles attempting to diffuse through the membrane. For example, the positively charged "perm-selective membranes" of Sollner [26] are made by the sorption of protamine cations on porous nitrocellulose, and the negatively charged ones are made by oxidation of the porous nitrocellulose. In the latter case there will be an equivalent number of sodium gegenions. Owing to the influence of the Donnan electrical effect, discussed in the next chapter, such an oxidized membrane is more permeable to sodium ions than to chloride ions because the concentration of sodium ions inside the pores is greater than the concen-

tration of the chloride ions, an excess always being present to balance the bound negative carboxyl ions. Hence in sufficiently dilute solution the membrane will behave as if it were permeable to sodium ions but almost impermeable to sodium chloride. Such membranes will produce nearly the maximum thermodynamically possible electrical potentials of concentration cells and may be used to determine the concentration or activity of ions in aqueous solutions.[27] Similarly Marshall [28] has shown that clay membranes may be used for determining the concentration of ions in aqueous and colloidal solutions. If such a membrane is interposed between 0.01 M potassium citrate and pure water, the transportation of water to the salt solution across the membrane, through electroosmosis, may be as much as 90 cc per square decimeter per hour.

Anomalous osmosis arises from differential ion permeability. It is a kinetic, not an equilibrium, condition. If the water flows from the dilute to the concentrated salt solutions at an accelerated rate the process is called positive anomalous osmosis. If the initial direction of flow is from the concentrated to the dilute solution it is called negative anomalous osmosis.

Membranes of living cells appear to be only one or two molecules thick, and they show a high electric resistance, which indicates comparatively poor permeability to ions. Nonprotein substances of molecular weight of 10,000 are too large to pass through the pores of a fertilization membrane of the egg of an echinoderm.[29]

Many instances are known to biologists where substances are *concentrated* by secretory or excretory organs or in certain cells, as compared with the surrounding fluid. Most of these cases have not received adequate explanation.

REFERENCES

Diffusion

Reviewed by WILLIAMS, J. W. and CADY, L. C., *Chem. Rev.*, 14, 171 (1934); DUCLAUX, J., *Diffusion dans les Liquids* and *Diffusion dans les Gels et les Solides*, Actualités Scientifiques et Industrielles, # 349 and 350, respectively, Hermann et Cie, Paris (1936); DEAN, R. B., *Chem. Rev.*, 43, 503 (1947)

Text References

1. KAWALKI, W., *Wied. Ann.*, 52, 166, 300 (1894); TUBANDT, C., REINHOLD, H. and JOST, W., *Z. anorg. Chem.*, 177, 253 (1928); SCHEFFER, J. D. R., *Z. physik. Chem.*, 2, 390 (1888)
2. CLACK, B. W., *Phil. Mag.*, 16(6), 863 (1908); TISELIUS, A. and GROSS, D., *Kolloid Z.*, 66, 11 (1934)

3. For graphs of the best available results for potassium chloride selected from International Critical Tables and Landolt-Börnstein Tabellen, *see* MCBAIN, J. W. and DAWSON, C. R., *Proc. Roy. Soc.*, **148A**, 33 (1935)

4. NORTHROP, J. H. and ANSON, M. L., *J. Gen. Physiol.*, **12**, 543 (1929). For further standardization and development of technique, see MCBAIN, J. W. and LIU, T. H., *J. Am. Chem. Soc.*, **53**, 59 (1931); DAWSON, C. R., *J. Am. Chem. Soc.*, **55**, 432 (1933), and lastly MCBAIN, J. W. and DAWSON, C. R., *loc. cit.*

5. MCBAIN, J. W. and LIU, T. H., *J. Am. Chem. Soc.*, **53**, 59 (1931); GORDON, A. R., *Annals N.Y. Acad. Sci.*, **46**, 285 (1945)

6. KING, C. V. and CATHCART, W. H., *J. Am. Chem. Soc.*, **58**, 1639 (1936)

7. GORDON, *loc. cit.*

8. HARNED, H. S., *Chem. Rev.*, **40**, 461 (1947)

9. LAMM, O. and POLSON, A., *Biochem. J.*, **30**, 528 (1936)

10. TISELIUS, A. and GROSS, D., *Kolloid Z.*, **66**, 11 (1934)

11. LAMM and POLSON, *loc. cit.*; GROSTING, L. J. and MORRIS, M. S., A.C.S. meeting, Chicago, 1948

12. COULSON, G. A., COX, J. T., OGSTON, A. G. and PHILPOT, J. ST. L., *Proc. Roy. Soc.*, **192A**, 382 (1948)

13. MCBAIN and LIU, *loc. cit.*

14. SUTHERLAND, W., *Australasian Assoc. Adv. Science*, **10**, 117 (1904); *Phil. Mag.*, **9**(6), 781 (1905); for the smallest molecules the numerical factor is 4 instead of 6.

15. EINSTEIN, A., *Ann. Physik*, **19**(4), 289 (1906); *Z. Elektrochem.*, **14**, 235 (1908)

16. NERNST, W., *Z. physik. Chem.*, **2**, 613 (1888); extended by Haskell for polyvalent electrolytes or charged colloid particles together with their compensating ions.

17. DEAN, R. B., *Chem. Rev.*, **41**, 503 (1947)

18. GORDON, *loc. cit.*

19. MCBAIN, M. E. L., *J. Am. Chem. Soc.*, **55**, 545 (1933)

20. ADAMSON, A. W., *J. Chem. Phys.*, **15**, 762 (1947)

21. VINOGRAD, J. R. and MCBAIN, J. W., *J. Am. Chem. Soc.*, **63**, 2008 (1941). On page 2011 N should read "moles" instead of "equivalents."

22. DEAN, R. B., *J. Am. Chem. Soc.*, **67**, 31 (1945)

23. MCBAIN, J. W. and DAWSON, C. R., *J. Am. Chem. Soc.*, **56**, 52 (1934)

24. LIESEGANG, R. H., *Chemische Reactionen in Gallerten*, Pamphlet (1896)

25. MORSE, H. W., *J. Phys. Chem.*, **34**, 1554 (1940)

26. ABRAMS, I. and SOLLNER, K., *J. Gen. Physiol.*, **26**, 369 (1943); GREGOR, H. P. and SOLLNER, K., *J. Phys. Chem.*, **50**, 88 (1946)

27. SOLLNER, K., *J. Am. Chem. Soc.*, **65**, 2260 (1943)

28. MARSHALL, C. E. and BERGMAN, W. E., *J. Am. Chem. Soc.*, **63**, 1911 (1941); *J. Phys. Chem.*, **46**, 52, 325 (1942); MARSHALL, C. E. and KRINBILL, C. A., *J. Am. Chem. Soc.*, **64**, 1814 (1942)

29. DEAN, R. B. and MOORE, A. R., *Biodynam.*, **6**, 159 (1947)

CHAPTER *15*

OSMOTIC PRESSURE

AND MEMBRANE EQUILIBRIUM

According to Einstein's dictum, every rigid particle in a solution should be equivalent to a single molecule and should therefore cause the same osmotic effects. Molar solutions of ideal particles or molecules should thus contain 60.6×10^{22} individuals per liter and give a lowering of freezing point of $1.858°$. Most colloidal solutions show nothing like this lowering, because they contain too few particles per liter. For example, a 0.5% gold sol with cubical particles 40 Å units in diameter is only 0.0000067 M, containing only 4×10^{18} particles per liter and with a predicted lowering due to these particles of $0.000012°$, which is far too small to be observed. Even the accompanying free ions compensating for the charge on each particle do not suffice to make this appreciable. Hence, most sols have vapor pressures, boiling points, and freezing points that differ measurably from that of water only by extraneous sources of error, such as impurities or dissolved gases.

A large group to be discussed in Chapters 17, 18, and 20 exhibits large osmotic effects, comparable with ordinary solutions.

Now osmotic pressure is a much more sensitive measure of the kinetic energy of the particles, since a molal solution has an osmotic pressure of about twenty-five atmospheres, and pressures may be measured in millimeters of water. The gold sol discussed in the first paragraph should exhibit an osmotic pressure due to the particles alone of 1.8 mm of water. The older scientific literature abounds in simple measurements of osmotic pressure, but their significance is questionable because of the influence of impurities and electrical effects to be discussed below. It would be necessary to show that these

results were actually equivalent to those obtained by thermodynamic methods and, for the most part, they do not refer to separate, uncharged particles.

The equipment employed is usually the osmometer, which separates the solution from the solvent by a membrane impermeable to the colloid. The solution is provided with a manometer with which its excess of hydrostatic pressure over the solvent is measured. The membrane may be a parchment tied over an inverted thistle tube, or a collodion sac, or an ultrafilter membrane. It is possible to work accurately with only 0.5 cc of solution.[1]

Of recent years, however, a number of carefully established measurements have been made [2] following upon Sörensen's study of egg albumin [3] and Adair's studies of hemoglobin.[4] For egg albumin the molecular weight found is 34,000, for hemoglobin 67,000, and for serum albumin 59,000 and 72,000. (Cf. Svedberg's ultracentrifuge values 40,500, 68,000, and 66,900, respectively.)

Since osmotic pressure is supposedly proportional to the number of particles present in solution, the value for any solution is dependent on the temperature, solvent used, pH, or any other factor which influences aggregation or dissociation. Meyer and Lühdemann [5] found that molecules of very long chains give by any of the physical chemical methods an "apparent" molecular weight tending in more concentrated solution towards 240, regardless of the known true molecular weight. Their explanation is that these long chain molecules are so flexible and loose-jointed that different parts of the same molecule are vibrating or moving almost independently. They found that with very dilute solutions this effect may be eliminated. Therefore it is important that the solution should be studied in extreme dilution, especially for nonaqueous or viscous solutions.

It has been commonly taken for granted that a solution containing as little as 1% by weight of a substance of extremely high molecular weight corresponds from the standpoint of physical chemistry to an almost infinite dilution of an ordinary substance of low molecular weight. This however appears not to be the case, and as is pointed out in the discussion by Kratky and Musil,[6] the deviations from ideal behavior are more nearly dependent on actual weight concentration than upon the number of molecules or particles in the unit of volume. The various ways in which these deviations have been interpreted and recorded are referred to in the article by Kratky and Musil.

Several authors, particularly Duclaux and Dobry,[7] had likewise observed that the osmotic pressure rises much more than proportion-

ately to the concentration. Adair [8] has added salt to protein solutions, producing large changes in the activity of the protein ion. He calculates the molecular weight from the usual van't Hoff law by correlating this change in activity with a nonideal distribution of salts across the membrane, rather than with the observed abnormal osmotic pressure.

Membrane or Donnan Equilibrium[9]

Wilhelm Ostwald in 1890 pointed out some interesting results that must occur in a system in which two solutions containing electrolytes are separated by a membrane which is permeable to all but one of the ions. Donnan in 1911 was the first to formulate these general phenomena, which are of universal occurrence where through any means an electrical charge is confined to a certain point or within a definite space. All other freely movable ions must adjust themselves accordingly. It has turned out that the principle involved is quantitatively the same as that developed by Gouy in 1910 with the diffuse double layer surrounding charged particles on surfaces, which again is the same as the ionic atmosphere derived by Debye and Hückel in 1923 for the space surrounding each ion or charge in an ordinary solution.

Although the general thermodynamic principle had been formulated by Gibbs in 1878 (Equation 78), it was Donnan who first applied the thermodynamic relations to electrically charged particles and ions. Donnan pointed out the many results which follow in respect to influence upon hydrostatic pressure in an osmometer, membrane potential, and hydrolysis. The effects are universal, although they hardly occupy the dominant position in colloid science ascribed to them by Jacques Loeb [10] in his monograph on proteins. The results are rigorous if expressed in terms of chemical potential or activities or of concentrations when each concentration is multiplied by its activity coefficient.

The principle for univalent ions may be most simply stated as follows. The product of the activities of any pair of oppositely charged, freely diffusable ions is constant throughout the system. The operation of this principle may be seen by consideration of the following diagrams.

Let the initial state of the system be a solution of the ions Na^+ and R^- of concentration c_1 or z, separated from a solution of sodium chloride of concentration c_2 by a membrane which is freely permeable to all but the R^- ions:

$$
\begin{array}{cc|cc}
Na^+ & R^- & Na^+ & Cl^- \\
c_1 & c_1 & c_2 & c_2 \\
z & z & &
\end{array}
$$
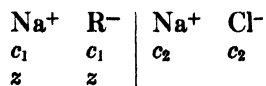
If now the vo█████ each side of the membrane are kept constant and equal, the di█████ions will rearrange themselves by sodium and chlorine ions passing through the membrane in equivalent quantities y, leaving the outer solution on the right hand side of concentration $c_2 - y$, or x. The letters are those most commonly used in discussions of this subject. The final equilibrium will then be:

$$
\begin{array}{ccc|cc}
\text{Na}^+ & \text{R}^- & \text{Cl}^- & \text{Na}^+ & \text{Cl}^- \\
c_1 + y & c_1 & y & c_2 - y & c_2 - y \\
y + z & z & y & x & x
\end{array}
$$
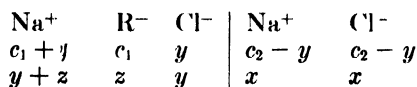
The Donnan principle requires that the product of the concentrations, or more strictly activities, of sodium and chlorine ions shall be the same on both sides at equilibrium. Hence

$$
\text{Na}_1^+ \cdot \text{Cl}_1^- = \text{Na}_2^+ \cdot \text{Cl}_2^- \text{ or } x^2 = y(y+z), \text{ or } \frac{\text{Na}_1^+}{\text{Na}_2^+} = \frac{\text{Cl}_2^-}{\text{Cl}_1^-}
$$

Since the same principle extends to all diffusible ions and since in aqueous solutions hydrogen and hydroxyl ions are always present, with ionic product K_w, we must write:

$$
\frac{\text{Na}_1^+}{\text{Na}_2^+} = \frac{\text{Cl}_2^-}{\text{Cl}_1^-} = \frac{(\text{OH})_2^-}{(\text{OH})_1^-} = \frac{\text{H}_1^+}{\text{H}_2^+}
$$

In other words, the membrane will always cause a certain small amount of hydrolysis.

It is evident that the chlorine ion (and an equal amount of sodium ion) has been partly hindered by the presence of the nondiffusible R^- ion from entering through the membrane. The extent to which this occurs is illustrated by the following table:

c_1	c_2	x/y
0.01	1	1.01
0.1	1	1.1
1	1	2
1	0.1	11
1	0.01	99

It is seen that the sodium chloride is almost quantitatively barred out if it is very dilute in comparison with the relative concentration of R^-. Only the ratios, not the absolute concentrations, matter. Equally important is the observation that the Donnan effect is almost negligible if the diffusible electrolyte is present in very high relative concentration. Much use is made of this latter principle in studies of diffusion, ultrafiltration, and ultracentrifuge, but care must be taken that the ratio be high enough.

If the salt, $Na^+ + R^-$ of original concentrat██ ███ exposed through the membrane to pure water, x of H^+ ion w██ ███se in producing x of Na^+ and x of OH^- outside, where $x = \sqrt{K_w \cdot c_1}$, c_1 being the concentration of the sodium ion inside at equilibrium. In dilute solutions this becomes of importance, and for weak acids or weak bases it is correspondingly magnified. Again it follows that if in dialysis the outside water is continually removed, complete hydrolysis will ultimately ensue.

If instead of using diffusible salt with a common ion we use a different salt, we have for the initial state:

$$\begin{array}{cc|cc} Na^+ & R^- & K^+ & Cl^- \\ c_1 & c_1 & c_2 & c_2 \end{array}$$

and for the final state:

$$\begin{array}{cccc|ccc} Na^+ & K^+ & R^- & Cl^- & Na^+ & K^+ & Cl^- \\ c_1 - z & x & c_1 & y & z & c_2 - x & c_2 - y \end{array}$$
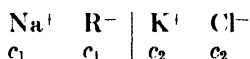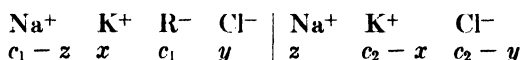
Two consequences arise from the Donnan equilibrium. The first is that the observed osmotic pressure, that is, the difference in hydrostatic pressure of the two sides at equilibrium, will not be that of R^- except for a very large excess of added diffusible salt. Thus for 10-fold excess the observed pressure will be still 4% too large. The osmotic pressures generally recorded in the literature often differ widely from these calculations.

The second consequence of the Donnan distribution is that a potential difference e_m is set up at the membrane, where

$$e_m = \frac{RT}{F} \cdot \ln \frac{c_{outside}}{c_{inside}}$$

or, more strictly,

$$e_m = \frac{RT}{nF} \cdot \ln \frac{c_2}{c_1}$$

where R is the gas constant, T the absolute temperature, F is 96,500 coulombs, and ln represents the natural logarithm ($= 2.3026 \log_{10}$). This has been verified within about 50%. Naturally no electromotive force is observed if reversible electrodes are placed at each side and the system is kept at equilibrium. It is necessary to measure the solutions separately; then their difference is equal to e_m.

It is then clear that the unequal distribution of ions is of universal occurrence whenever any charges are fixed in position. The effects have to be taken into account unless they are suppressed by the presence of a large excess of a diffusible salt.

Application to Proteins

Before discussing the Donnan equilibrium for protein solutions, tribute must be given to the invaluable services of Jacques Loeb in impressing upon all concerned the predominant role of hydrogen and hydroxyl ions in the behavior of amphoteric or acidic and basic colloids. In previous work the value of the acidity or pH had been left mainly to chance, and the results were often as confusing or uncertain as is the study of an unstable colloid without reference to its stabilizing agents. It need only be mentioned here that the importance of pH is unaffected by whether the reaction of a protein with an acid is wholly chemical, sorption, or both. Likewise, as is emphasized in Chapter 9, even when pH is fixed, the influence of various other ions is clearly shown in their lyotropic behavior which affects even the position of the isoelectric point. As has been seen, the influence of pH is of very great importance even in the electrokinetic movement of air bubbles where no chemical groups come into consideration.

At first sight it is not evident how the Donnan equilibrium can play any role with so weak an acid or base, and there is certainly none at the isoelectric point. However, if within the membrane we again consider, for simplicity's sake, a weak monovalent base B_0 of large size and of dissociation constant K_b, present at initial concentration c_0, and place outside the membrane an aqueous solution of hydrochloric acid, the final equilibrium will be:

$$
\begin{array}{ccccc|cc}
B_0 & B^+ & H^+ & Cl^- & & H^+ & Cl^- \\
c_0 - z & z & y & y + z & & x & x
\end{array}
$$
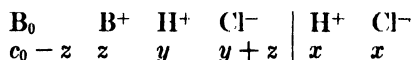
and we may write the equations:

$$x^2 = y(y + z) \quad \text{and} \quad z = \frac{c_0 y}{\dfrac{K_w}{K_b} + y}$$

The osmotic pressure against water would initially be RTc_0, where $x = 0$. This would be increased as acid is added by the amount

$$\frac{RT(x - y)^2}{y}$$

which passes through a maximum and for very large values of x again asymptotically falls to zero. (For the actually polyvalent proteins the final osmotic pressure remains as many fold that of the isoelectric protein.) Therefore, with change of pH on the acid side of the isoelectric point, osmotic pressure passes through a maximum as is shown in Figure 15·1. Likewise, there would be a similar maximum on the alka-

line side since protein is amphoteric. The effect of all neutral salts would be to depress these maxima.

Quoting from Loeb, who regarded Donnan effects as the essence of colloid behavior,[11]

> The question may be asked: Why are proteins termed colloids? The answer is that proteins show colloidal behavior only in regard to the influence of electrolytes on four well-defined properties; namely, membrane potentials, osmotic pressure, swelling, and that form of viscosity which is due to the swelling of submicroscopic particles. This influence of electrolytes is similar on all four properties.

In Loeb's figure (see our Fig. 15·1) one upper group occurs in diminishing order: nitric, trichloracetic, hydrochloric, phosphoric, and oxalic, which here are all effectively monovalent and have similar, although by no means identical, action. Very different and much less is the effect of the strong, divalent sulfuric acid. The right end of the curve is approximately at the isoelectric point, and it will be noted that none of the four properties listed falls to zero but all pass through a real minimum.

In general, the Donnan effects are real, but often only minor. For
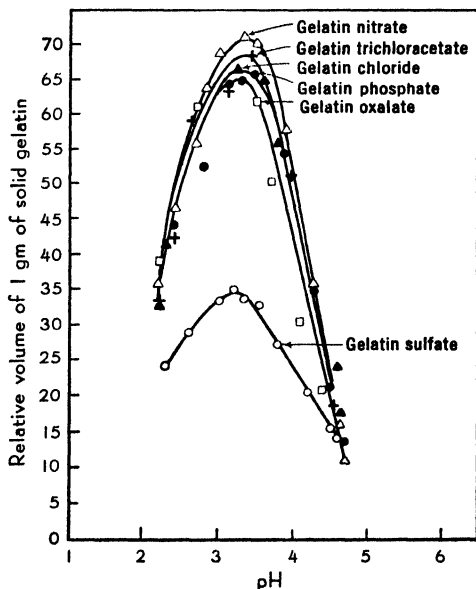
FIGURE 15·1 Influence of HCl, HNO$_3$, H$_3$PO$_4$, H$_2$SO$_4$, trichloracetic and oxalic acids on the swelling of gelatin. Abscissas are the pH, ordinates the volume of gelatin. The curves for all the acids are practically identical except that for H$_2$SO$_4$, which is about one half as high as the curves for the other acids. Similar diagrams were obtained where the ordinates were, instead, the viscosity of gelatin sol, the apparent osmotic pressure of gelatin sol, or the swelling of gelatin jelly. Taken from Loeb, Text Ref. No. 10.

example, the effect of *salt* in *increasing* the viscosity of a solution of the soap sodium palmitate is ten times greater than the largest effect recorded of any acid on gelatin.

Other Applications

The influence of the Donnan unequal distribution of charges has always to be remembered. We have seen that every ion or charge, whether free in solution or fixed to a surface or colloidal particle, is surrounded by such an ionic atmosphere. Hence in ultrafiltration, for example, we have the effects discussed in Chapter 8, not only the Donnan distribution of ions between solution above and below the membrane when time is allowed for its establishment, but also the influence of rate of filtration and the difficulty of filtering very dilute electrolytes through a charged membrane. Through it an ion may be accumulated against its own concentration gradient by making use of an ion-selective membrane such as a positive protamine-collodion membrane.

REFERENCES

Text References

1. For references see PAULI, WO. and FENT, P., *Kolloid-Z.*, **67**, 289 (1934)
2. PAULI and FENT, *loc. cit.*
3. SÖRENSEN, S. P. L., *Compt. rend. trav. lab. Carlsberg*, **12**, 262 (1915-17)
4. ADAIR, C. S., *Proc. Roy. Soc.*, **109A**, 292 (1925)
5. MEYER, K. H. and LÜHDEMANN, R., *Helv. Chim. Acta*, **18**, 307 (1935)
6. KRATKY, O. and MUSIL, A., *Z. Elektrochem.*, **43**, 326 (1937)
7. DOBRY, A., *J. chim. Phys.*, **32**, 46, 50 (1935); DUCLAUX, J., *Compt. rend.*, **153**, 1580 (1911)
8. ADAIR, G. S., *Trans. Farad. Soc.*, **31**, 98 (1935)
9. DONNAN, F. G., *Z. Elektrochem.*, **17**, 572 (1911); *Chem. Rev.*, **1**, 73-90 (1924); *Trans. Farad. Soc.*, **31**, 80 (1935); *Z. physik. Chem.*, **168**, 369 (1934); DONNAN, P. G. and GUGGENHEIM, E. A., *Z. physik. Chem.*, **162**, 346 (1932); BOLAM, T. R., *The Donnan Equilibrium*, G. Bell and Sons, Ltd., London (1932); LOEB, J., *Science*, **56**, 731 (1922); ADAIR, G. S., *Trans. Farad. Soc.*, **33**, 1106 (1937)
10. LOEB, JACQUES, *Proteins and the Theory of Colloidal Behavior*, McGraw-Hill Book Co., Inc., New York (1922)
11. BOGUE, R. H., *Theory and Application of Colloidal Behavior*, McGraw-Hill Book Co., Inc., New York (1924) p. 39

ULTRACENTRIFUGES AND CENTRIFUGES

If a handful of sand is dropped into a glass cylinder full of water and the grains of sand are all of uniform size, they will settle with uniform speed, and there will be one sharp boundary between the uppermost sand grains and the clear water left behind on top. Obviously, it is possible to calculate the effective size of the sand grains from the rate at which they fall, by applying Stokes' law. Now if there were present a proportion of particles of another much smaller size, two upper boundaries would be observed corresponding to the two rates; that is, rate of sedimentation discriminates sharply between particles of different sizes. If there is only one rate or sharp boundary, the particles are monodisperse, that is, all of the same size.

Among the important tools of colloid science are sedimentation, centrifugation, and ultracentrifugation. Sedimentation takes place under gravity. The effective force is greatly magnified by using a centrifuge which can readily increase the field from a thousand to a million times gravity. An ultracentrifuge is not necessarily a high speed centrifuge, but it is one in which convection is avoided so that undisturbed sedimentation takes place in the centrifugal field. Then there must be some means of investigation of the distribution of the material in the ultracentrifuge cell either during centrifugation or immediately thereafter. An ultracentrifuge cell is always very small — Svedberg's cells are only 0.8–1.5 cm long if completely filled; it has proved impossible to avoid convection in larger volumes of liquid.

Sedimentation under gravity is very important in soil science and has been practiced since the beginning of the 19th century. It is, however, useful only for particles larger than about 0.1 micron. Extraordinary precautions are required to avoid convection in colloidal sols sufficiently that their sedimentation under the influence of gravity may be observed, as was discussed in Chapter 7.

In the absence of convection there are only two forces operating upon the suspended particles, whether these are of colloidal or molecular dimensions. The first is sedimentation, which depends upon the relation between the density of the particle and the buoyancy of the medium and is inversely proportional to the friction f, which opposes the motion of a particle. The other is diffusion, which immediately sets in whenever there is any difference in concentration and which is proportional to the concentration gradient for any given point. A mathematical analysis has been given by Mason and Weaver [1] based simply upon Stokes' law for sedimentation and Fick's law for diffusion. A corresponding analysis has been carried out by Archibald [2] for the centrifugal field with fairly similar results.

The centrifugal field is not uniform, as contrasted with the gravitational field, but this is easily allowed for in the calculation since at any point the centrifugal force equals

$$\omega^2 x$$

where ω is the rotational speed in radians per second, and x is the distance from the axis of rotation. $\omega = 2\pi$ times the number of revolutions per second (r.p.s.); hence the centrifugal force per centimeter of radius is

$$0.0404 \text{ r.p.s.}^2 g$$

where g is 980 dynes, the acceleration due to gravity.

The buoyancy effect for a particle which merely displaces an equal amount of liquid is simply the difference between the density of the particle and the density of the liquid, but the general expression embracing all possible gases, particles, molecules, and ions is

$$1 - \bar{V}\rho$$

where ρ is the density of the solution and $\bar{V}$ is the partial specific volume of the colloid or solution. The partial specific volume $\bar{V}$ is easily calculated from a determination of densities, by the methods given in "Thermodynamics" by Lewis and Randall, Chapter 4, and in particular the method of graphic intercepts shown there in detail.

It was a great advance in the understanding of colloids, particularly the stable colloids such as proteins, when Svedberg and Nichols in 1923 invented the ultracentrifuge. This was followed by many later improvements in Svedberg's laboratory at Upsala, and in those of Beams at Virginia, and of the author at Stanford University. For this great achievement Svedberg was awarded the Nobel prize.

Rate of Sedimentation

There are two primary measurements to be made with the ultra-centrifuge. The first, and by far the most usual, is rate of sedimentation. The observation consists of finding the positions x_1, x_2, etc., the boundary corresponding to a particular species, at times t_1, t_2, etc., which directly gives the sedimentation velocity dx/dt at a known value of $\omega^2 x$. The sedimentation constant s is then the linear velocity referred to a field of only 1 dyne for any given species and is usually, by a rather arbitrary calculation (see below), referred to water at 20°C for solvent:

$$s = \frac{dx/dt}{\omega^2 x}$$

$$s_{20°} = \frac{dx/dt}{\omega^2 x} \cdot \frac{\eta(1 - \bar{V}\rho_0)}{\eta_0(1 - \bar{V}\rho)} \text{ cm per sec}$$

where η and ρ are the viscosity and density of the solution, and η_0 and ρ_0 those of pure water at 20°C.

For a particle of weight m, the centrifugal force, corrected for buoyancy, acting on it may be equated with the frictional force f which it encounters. Integrating and indicating logarithms to the base e by the symbol ln:

$$m = \frac{f \cdot \ln x_2/x_1}{\omega^2 x \cdot (1 - \bar{V}\rho) \cdot (t_2 - t_1)}$$

This assumes that no gegenions have to move with it; so a buffer, preferably of potassium chloride, is employed to minimize electrical effects. (See Chap. 14.) Now if measurements of diffusion *in the same medium* are also carried out:

$$D = \frac{RT}{N} \cdot \frac{1}{f}$$

Hence when this is substituted in the equation above we derive the particle weight or molecular weight M in grams

$$M = \frac{RT \ln x_2/x_1}{D \cdot \omega^2 x \cdot (1 - \bar{V}\rho) \cdot (t_2 - t_1)}$$

or

$$M = \frac{RTs}{D(1 - \bar{V}\rho)}$$

This equation gives no information as to the shape of the particle.

In case it is justifiable to assume that the particles are spherical as well as uncharged and their density or radius can be estimated, the diffusion constant can be calculated instead from the Sutherland-Einstein equation (Chap. 14):

$$D = \frac{RT}{N} \cdot \frac{1}{6\pi\eta r}$$

The greatest value of measuring rate of sedimentation is the decision as to whether a colloid is monodisperse or polydisperse. In the latter case there is a corresponding number of boundaries or they may blend into a wide band. Since diffusion also blurs the boundary, the experiment is performed as quickly as possible by using a high centrifugal force. This may require for a protein one or more hours at one or more hundred thousand times gravity.

Sedimentation Equilibrium

The other important method of study with the ultracentrifuge is the observation of the equilibrium or steady state that is brought about when the rate of sedimentation is kept so low that it is largely offset by the rate of diffusion back again. In the steady state the distribution of any given size of particle or molecule is logarithmic, just as in the earth's atmosphere. Such measurements are carried out with a low centrifugal force, but necessarily allowing much time for diffusion to produce the equilibrium distribution. For example, for typical proteins many days are required at a centrifugal force of only five thousand times gravity. This is the only thermodynamically established method, and it gives no information as to the shape of the particles. Archibald [3] has shown how the approximate position of the equilibrium may be calculated from the earlier stages without waiting for it to be actually attained.

To obtain the quantitative expression we use the thermodynamic equation of Lewis and Randall [4] for equilibrium:

$$\Delta \bar{F} = 0 = \frac{\delta \bar{F}}{\delta x} \cdot dx + \frac{\delta \bar{F}}{\delta c} \cdot dc + \frac{\delta \bar{F}}{\delta p} \cdot dp$$

Now it has been customary to assume for the sake of simplicity that the activity coefficient is unity and that one is dealing with perfect, uncharged nonelectrolytes. Substituting in this equation the expressions relating partial molar free energy $\bar{F}$ to position x, concentration c, and pressure p, we arrive at the equation:

$$M = \frac{2RT \ln c_2/c_1}{(1 - \bar{V}\rho)\omega^2(x_2^2 - x_1^2)}$$

The same equation must follow from setting the sedimentation velocity equal to the rate of diffusion.

After a discussion of the two main kinds of measurement which can be performed with the ultracentrifuge, it now remains to describe the instruments in which they have been realized. It should be recalled that the essential condition is that the sedimentation must not be disturbed by convection. Then there must be the method of examining the distribution of the contents of the cell.

The Oil Turbine Centrifuges of Svedberg [5]

The Svedberg ultracentrifuge for measuring *sedimentation velocity* is a very impressive apparatus, which requires a couple of rooms to be housed. The Svedberg ultracentrifuge has enabled fields up to 400,000 times gravity to be regularly used. This subjects the available constructional materials to stresses near their limits.

The small cell consists of three quartz or fused silica discs cemented or pressed together, the center one having cut in it a wide slit whose sides are sectorial, that is, radiating from the center of rotation, as

FIGURE 16-1 Diagram showing center disc of cell, with sectorial sides; that is, radiating from the center of rotation.
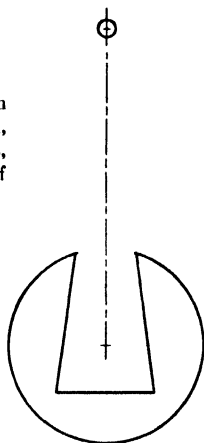
shown in Figure 16·1. The two outer discs provide the other two plane sides of the cell. The cell with its housing is placed in a large rotor whose axes carry the oil turbine blades, driven by a 24 h.p. motor. The rotor is spun in hydrogen at a few millimeters pressure. The transparent cell allows filtered light, visible or ultraviolet, to pass through each time the cell traverses the light beam and permits a photographic record of the absorption or refraction of the light, from which the concentration of the contents of the cell at that point is deduced. Typical

Distance from meniscus, cm.

photographs are shown in Figures 16·2 and 16·3. As is shown in Figure 16·2, the absorption of light in the different portions of the cell at successive periods is recorded.

The criterion for such a photograph is that the solution is uniform below each boundary unless a further boundary is also present. The solution below the boundary becomes more dilute with time owing to the sectorial shape of the cell and the increasing rate of sedimentation at greater radii. Similarly the boundary travels faster and faster, and

Meniscus

Protein components

Index
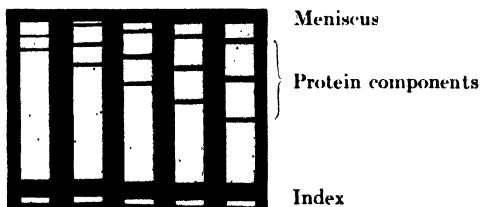
FIGURE 16·3 Sedimentation pictures for *Limulus* hemocyanin, obtained by the Toepler schlieren method at pH 6.8, showing the three fastest sedimenting components of s, 56.5×10^{-13}, 34.6×10^{-13}, and 16.1×10^{-13}. Centrifugal force, 120,000 times gravity. Time between exposures, 5 minutes. Taken from Eriksson-Quensel and Svedberg, *loc. cit.*

it becomes more diffuse owing to the longer period of diffusion. In Figure 16·3 the three boundaries show the presence of three constituents. This is the same *Limulus* hemocyanin, using the Toepler schlieren method based upon refractive index.

There are many other methods of optical recording, such as that already referred to in connection with the electrophoresis cell, as illustrated in Chapter 13.

Much simpler is the Svedberg machine for measuring *sedimentation equilibrium*.[6] The low speed machine is mounted on ball bearings and is directly driven by an electric motor. It ranges from 500 to 15,000 g. A good example of the measurement of sedimentation equilibrium in a monodisperse system is given in Figure 16·4.
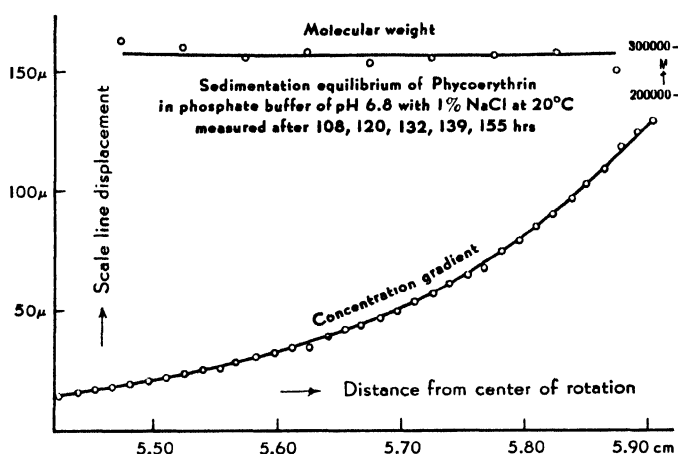
FIGURE 16·4 Relation between molecular weight and distance from center of rotation for phycoerythrin (M 290,000) at pH 6.8. Taken from Eriksson-Quensel and Svedberg, *loc. cit.*

The McBain-O'Sullivan Transparent Ultracentrifuge

Henriot and Huguenard invented the air-spun rotor.[7] This is shown in cross section in Figure 16·5. Compressed air at from 10 to 100 lb/sq in. pressure from the wind chamber *A* passes through the diagonal holes in the cone *B* and escapes between that cone and the rotor *C*. The rotor is spun and is sucked into the cone by the Bernoulli effect but is separated from it by the film of escaping air. When properly made, the spinning is so smooth that no lateral or vertical motion can be detected with the highest power microscope. The only limit to the speeds attained is the tensile strength of the material of the rotor, so that with a 2 cm hollow rotor a force of a million times gravity is attainable. MacHattie [8] has spun a solid 2.3 mm steel ball electrically at fifty-eight million times gravity although only one ball survived this test, the stress at the center of the ball being 320,000 lb/sq in. Beams, with a 0.5 mm solid ball has attained 430 million times gravity. Such contrivances have no usefulness for ultracentrifugation.

The Henriot and Huguenard arrangement shown in Figure 16·5 is not an ultracentrifuge because there is no prevention of convection.

However, it has been effectively used in a series of studies of virus by French and Belgian investigators.

McBain and O'Sullivan developed this into a transparent ultracentrifuge, with which in April, 1935, the first completely successful measurements, other than in a Svedberg ultracentrifuge, were achieved.[9] Figure 16·6 reproduces some results obtained with this ultracentrifuge. Although this instrument has not been used elsewhere, it is still the only completely thermostated ultracentrifuge because the rotor immediately acquires exactly the temperature of the slip stream of air and the air is drawn at constant pressure through a thermostat of any desired temperature. The rotor is customarily run at several thousand rps, and it has the other very great advantage that it takes only a few seconds to start or stop, so that, for example, living protoplasm would need be exposed to the ultracentrifugal field for only a few minutes.

FIGURE 16·5 The Henriot and Huguenard air-spun rotor and base in cross section.

In the McBain-O'Sullivan transparent ultracentrifuge the rotor is directly air driven, and the cell is placed in the rotor. The light comes up through the central air space, and a totally reflecting prism or periscope is held stationary inside the bottom cone of the rotor, as shown in Figure 16·7, to offset the light and pass it out through the cell. Because of the smaller dimensions of the cell, an ordinary long bellows camera can be used without parallax. Strictly monochromatic ultraviolet or visible light is provided through a prism and slit. For hemoglobin $s_{20} \times 10^{13}$ was found equal to 4.67, as compared with the best Upsala value of 4.63. At the lower limit of particle weight for this instrument was egg albumin, giving 3.64 instead of Svedberg's value of 3.55. The opaque or analytical ultracentrifuge of McBain and Leyda, described below, is much more accurate and powerful for monodisperse systems and may be used with accuracy even for ordinary chemical molecules.

The small size of the cell may raise a question, but it follows from the calculations by Mason and Weaver, confirmed by Svedberg's collaborators, that the resolving power of an ultracentrifuge is given by the expression

$$\omega^2 x h$$

where h is the height of the column of solution in the cell. Hence, if this product is kept constant the results must be identically the same

FIGURE 16·6 Some results obtained with the McBain-O'Sullivan transparent ultra-centrifuge. A is earthworm blood; B is egg albumin; C is edestin; D is beef hemo-globin; and E is Lansing poliomyelitis virus. A is taken from McBain and O'Sullivan, *J. Am. Chem. Soc.*, **57**, 780 (1935); B, C, and D are taken from Lewis, A. H., Doctorate Thesis, Stanford University (1939); E is taken from Loring, H. S., and Schwerdt, C. E., *Proc. Soc. Exp. Biol. Med.*, **62**, 289 (1946).
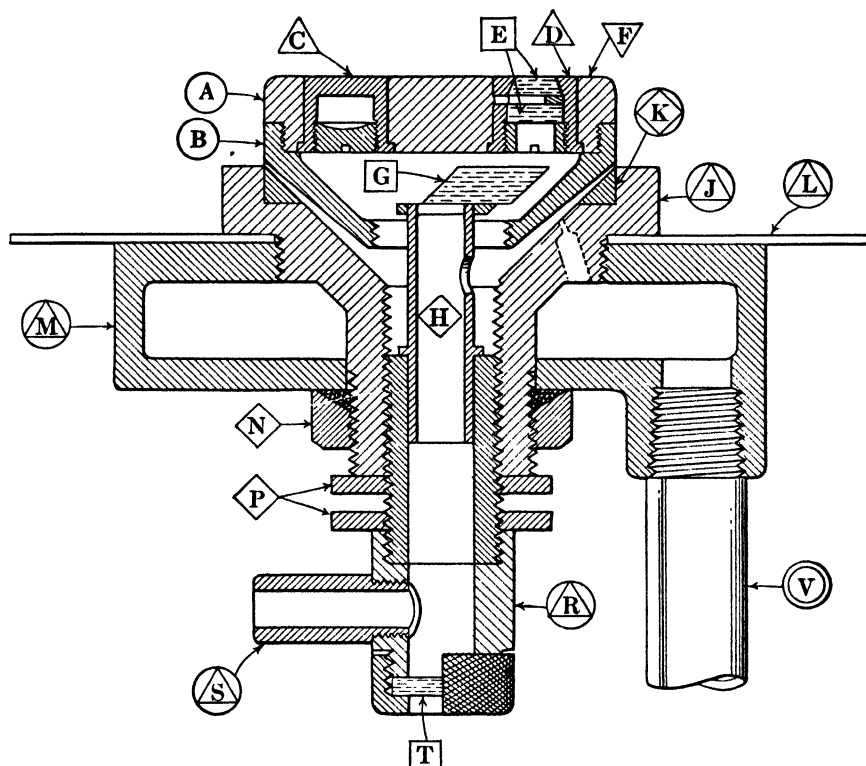
FIGURE 16·7 Details of rotor, stator, and periscope of the McBain-O'Sullivan air-driven transparent ultracentrifuge. A, cell-holder; B, rotor cone; C, dummy cell; D, cell; E, quartz discs; F, hard-rubber separator; G, periscope prism; H, periscope tube; J, stator; K, bronze insert; L, disc to which baffle strips are fastened; M, manifold; N, packing nut; P, lock-nuts; R, holder for central air inlet and window; S, connection to central air inlet; T, window; V, high pressure supply tube. The symbols indicate the materials of construction: O, Republic Steel Corporation's UMA 4, hardened and ground; △, Monel metal; □, quartz; ▽, hard rubber; ◇, brass; O△, duralumin (17ST); O◇, bronze; Ⓞ, aluminum.

for all sizes of cell. For particles of double molecular weight the centrifugal field should be only half as great to produce an identical pattern.

The directly air-driven ultracentrifuge of McBain and O'Sullivan is the only one that holds out hope of examining living protoplasm since protoplasm does not endure prolonged centrifuging, and this ultracentrifuge can be started and stopped within a minute. A sufficiently large living cell, such as that of *Chara cerotophylla*, or the egg of the salamander (which has given promising results), or possibly a slime mold, self-clarified by growth through filter paper, are the most likely to give results, though they must be biologically tested before

and after to be convincing. No existing ultracentrifuge is powerful enough to examine small living cells.

The Beams-Pickels Air-Driven Transparent Ultracentrifuges

Biscoe, Pickels, and Wyckoff [10] separated the air turbine drive from the rotor, which is suspended below it in an evacuated chamber by a spindle of piano wire which runs through a vacuum-tight oil seal. The rotor is similar to that of Svedberg and can develop a centrifugal force of 250,000 times gravity. Although far more expensive than the McBain-O'Sullivan type, it costs but a small fraction of that of Svedberg and it is the most commonly used transparent ultracentrifuge.

Capillary Tube Ultracentrifuges

It is found that capillary tubes inhibit convection, and although the sides are cylindrical instead of sectorial or radial, any flow occasioned is confined to an ultramicroscopic distance from the walls. Hence, within experimental error the results obtained can be identical with those of Svedberg's ultracentrifuge. Such capillary tube ultracentrifuges are so inexpensive as to be within the reach of any laboratory.

The "simplest ultracentrifuge with photographic recording," described by McBain and Lewis,[11] consists of a solid rotor carrying an ordinary capillary tube of approximately 1 mm bore. With it erythrocruorin of earthworm blood gave $s_{20} \times 10^{13} = 59.3$ as compared with the Upsala value of 60.9 and 4.72 for carboxyhemoglobin as compared with Steinhardt's Upsala value of 4.63.

The capillary tube ultracentrifuges can take many forms [12] as developed by Elford, Ford, Wilson, and the author. In a typical form, as developed by the last three mentioned (unpublished), thin-walled capillary tubes, 1 or 2 cm long and of less than 1 mm diameter, sometimes cylindrical and sometimes of elliptical cross section and sealed at one end, are placed in a ring A provided with radial holes on a directly air-driven rotor B, as in Figures 16·8 and 16·9, or laid across the rotor in radial grooves. The rotor consists of three parts, as shown in Figure 16·8 and photographed in Figure 16·9. There is the usual conical base with a hollow for the radiately perforated ring A, which holds the capillary tube and a cover screwed on to prevent the ring from flying off during centrifuging. It is found possible to stop the rotor after an interval, take out the tubes, place them on a photographic

enlarger, and obtain a highly magnified photograph which is easily measured. By interposing a slit, all but the central part of the tube may be masked, greatly improving the photograph. Absorption of light
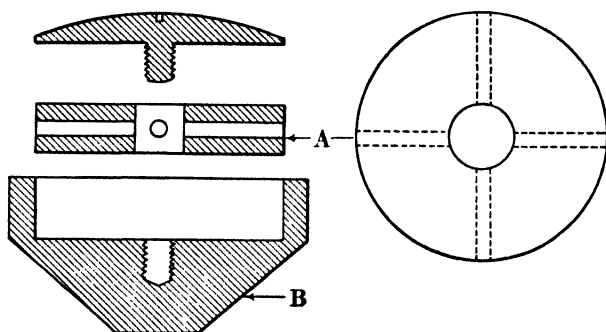
FIGURE 16·8 Diagram of a typical form of capillary tube ultracentrifuge. A is cross section (left) and plane view (right) of ring provided with radial holes for the insertion of tubes; B is the air-driven rotor.

may be used as in the experiments by Ford with clay, or refractive index as done by Wilson and by Ford, or opalescence as used by Elford. Figures 16·10 and 16·11 show results obtained by Ford using absorption of light, schlieren, and scattered light, respectively, in these tiny tubes.

The device used by McBain and Van Acker for determining the boundary or boundaries when studying small fragments of protein molecules of less than 2000 molecular weight is to insert the tube through a hole in the thin wall of an opaque box with the sealed end on the outside serving as a lens for a carbon arc lamp. When photographed at right angles, the opalescence and any variation of intensity
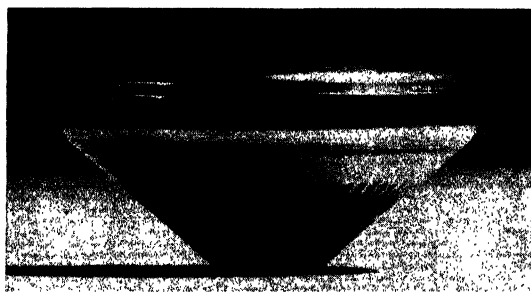
FIGURE 16·9 Photograph of the rotor B diagramed in Figure 16·8, closed, ready for use.
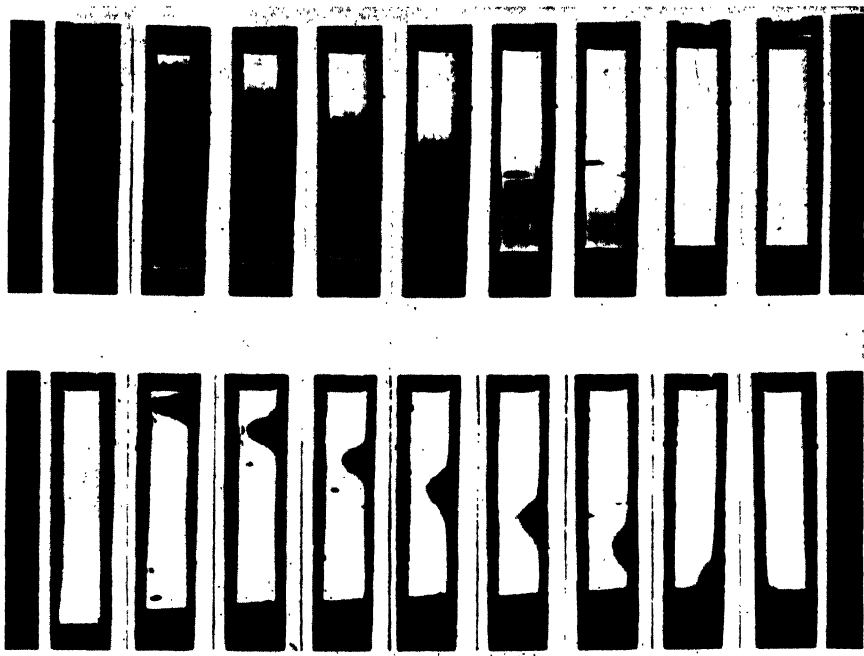
FIGURE 16·10 Results with earthworm blood obtained by Ford (not yet published), using absorption of light and schlieren, respectively. 800 rps, 4-minute intervals.

FIGURE 16·11 Results with Jersey skim milk, one day old, obtained by Ford (not yet published), using scattered light. 400 rps, centrifugal force about 8000 times gravity.

below the boundary are very apparent from the photograph. All of these directly air-driven ultracentrifuges have the advantage that they can be run at any exactly controlled temperature above, below, or at ordinary temperatures. They may also be used to determine rate of equilibrium.
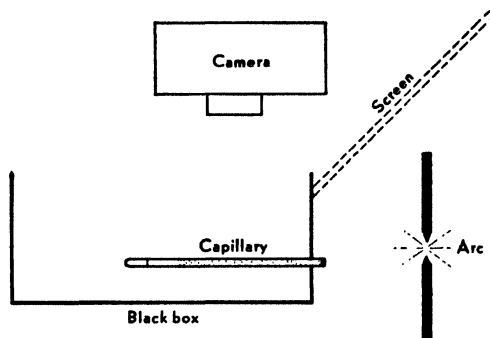
FIGURE 16·12 Device of McBain and Van Acker for photographing scattered ultraviolet light so as to identify the portion of sedimenting boundaries.

OPAQUE ULTRACENTRIFUGES FOR DIRECT ANALYSIS

The most accurate ultracentrifuge for monodisperse systems, and the one that has been used for ordinary molecules such as sucrose, is the opaque type of ultracentrifuge developed by the author at Stanford University. The important first step was to devise a two-piece rotor with a tight seal such as is shown in cross section in Figure 16·13, omitting C for the moment. Such a rotor is not an ultracentrifuge but has to be converted into one. A relatively thin-walled cover A screws into a thick-walled, air-driven rotor base B, a tight seal being provided by the film of pliofilm or other plastic D for aqueous solutions (Cellophane for petroleum derivatives) against a metal disc E provided with a loose slip ring F. When this box is screwed up finger-tight, it makes a seal that will hold a liquid hydrostatic pressure at any speed.

Now it is turned into an ultracentrifuge by placing in position any one of a number of inserts such as the McBain-Leyda insert C of Figure 16·13. Figure 16·14 is a photograph of A, B, and C. The McBain-Leyda insert consists of a set of stainless steel or platinum-iridium discs 0.003″ thick, 50 of smaller diameter and 50 of larger diameter being used, placed alternatively. Thus a horizontal annular space 0.076 mm deep is provided between the wider discs which completely avoids convection even when the rotor is gently stopped. Hence sedimentation takes place undisturbed by convection between the discs, whereas

the solution is uniform in the space outside the wider discs between them and the box and is of identical composition with that at the outside radius of the discs. A hole 1 mm square is cut on opposite sides of the top disc or insert and the outside homogeneous liquid is withdrawn for direct analysis. Any kind of chemical, physical, or biological analysis may be employed.
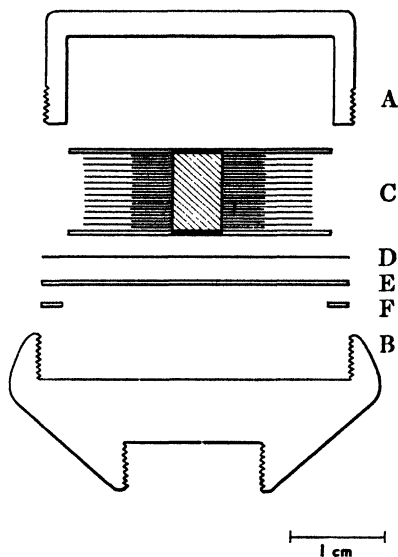
FIGURE 16·13 Cross section of opaque ultracentrifuge rotor with tight seal. A, the thin-walled cover; B, the air-driven rotor base; C, the McBain-Leyda insert; D, film of plastic; E, metal disc provided with slip-ring, F.

The advantage of this method is that it measures directly the actual amount of sedimentation in a given time in a given centrifugal field at a given distance from the center of rotation, namely, the radius of the larger discs. Diffusion does not enter in unless the ultracentrifuge is run too long. The sedimentation velocity of isoelectric egg albumin so determined is $s_{20} \times 10^{13} = 3.54$ as compared with Svedberg's value of 3.55. The *rate* of sedimentation of sucrose has been frequently used in the Stanford laboratories as a check on the correct running of this ultracentrifuge.

When this ultracentrifuge is allowed to run for a sufficient length of time, if the material therein is monodisperse, the sedimentation *equilibrium* is readily calculated from the analysis. However, another form of insert was used by McBain and Alvarez-Tostado [13] in which they obtained for the molecular weight of sucrose 331, 339, 353;
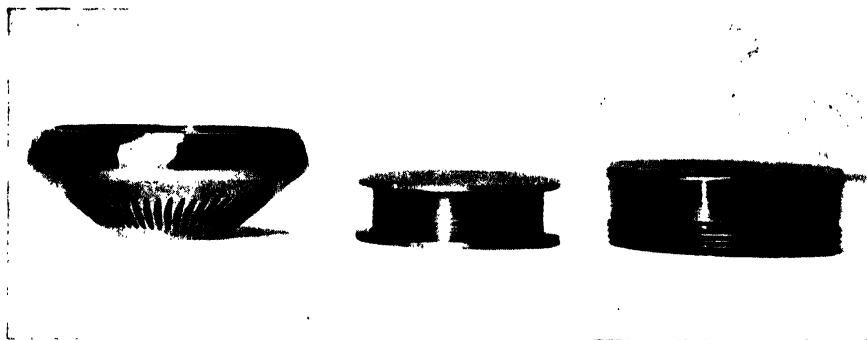
FIGURE 16-14 Photograph of A, B, and C of Figure 16-13.

mean 341, as compared with the known value of 342. Instead of having the discs in the center they used annular shaped washers of two widths all touching the walls of the box and therefore open at the center.

Svedberg has shown that it is possible to put a porous partition halfway down his transparent cell and then use an analytical method.

The appropriate methods of calculation for these and the various other ultracentrifuge methods have been summarized.[14]

Analytical centrifuges are not as satisfactory for polydisperse particles of a given substance as the transparent forms because then they have to be run at several speeds. Their field is where one substance is monodisperse, and this is not affected by the presence of other materials.

Since colloidal solutions containing anisotropic or elongated particles that may orient in the ultracentrifuge cannot be measured with certainty by optical means, soaps being a good example, and since in certain soap solutions some of the colloidal particles sediment outward while others rise toward the center, the only sure method of investigation is to use a nest of concentric rings as insert. Five such rings are pierced by hundreds of radial holes and after centrifugation each ring is pulled out separately and its contents analyzed. In this way the presence of the two groups of colloidal particles, sedimenting and creaming, was verified (unpublished) confirming the indications that had been given in the experiments of McBain and McBain in Svedberg's laboratory.[15]

McBain and Stuewer [16] showed that the simple Henriot and Huguenard top (Fig. 16-5) serves as a true ultracentrifuge for any immobilized system such as a jelly or a gel where convection is avoided.

Uses of Ultracentrifuges (Whether Equilibrium or Rate)

1. Determination of particle size, or molecular weight if there is only one molecule to each particle.

2. Determination as to whether the particles are monodisperse or polydisperse.

3. Determination of densities of the particle by changing the density of the medium until sedimentation changes to creaming. McIntosh and Selbie [17] found the density of *B. subtilis* spores to be 1.46, whereas phages and *Micrococcus (Staphylococcus) aureus* were 1.25.

4. Determination of the composition of the sedimenting particles by establishing by direct analysis the identity or nonidentity of the movements of their constituents.

5. Determination of the swelling pressure of jellies.

6. Determination of water bound by the particles.[18] Lansing and Kraemer [19] obtained an identical result for the system discussed by McBain. As regards jellies, it is significant that sedimentation of small molecules through them is unhindered, thus supporting the concept of the brush heap structure.

Svedberg's Fundamental Results for Proteins

The first results with Svedberg's ultracentrifuge were obtained with gold sols which, like almost all other colloidal preparations, proved to be polydisperse, although the range of sizes and the proportion of each size could be determined. Next Svedberg found the totally unexpected result that many proteins as supplied in nature are monodisperse in spite of their enormous particle size as compared with previously studied molecules.

Svedberg's monumental work is summarized in Table 16, which gives the sedimentation constant s_{20} multiplied by 10^{13}, the diffusion coefficient D_{20} multiplied by 10^{7}, the molecular weight M_s computed from sedimentation velocity and diffusion, the molecular weight M_e computed from sedimentation equilibrium, M_{calc} based upon the original suggestion of Svedberg (but hardly borne out by Table 16) that the molecular weights of proteins are multiples of 35,200, and f/f_o the ratio of the experimentally determined molecular frictional constant to that calculated for a spherical particle of the same mass. The departure from unity can be due to several factors, such as elongation, disc-shape, or other geometrical departure from spherical. Also included in f/f_o is the hydration of the proteins which is always present.

For example, for hemoglobin where the ratio is 1.25 Adair and Adair [20] showed that after taking hydration into account the ratio is reduced to 1.10. More information as to shape is afforded by studies of viscosity, for which many differing formulas have been proposed. One calculation represents egg albumin as an ellipsoid 2.3 times as long as it is wide.[21]

The table also gives the pH of the isoelectric point and the rate of change of electrokinetic mobility, the change in pH near the isoelectric point, du/dpH$_0$, the value being multiplied by 10^5.

It is seen from an examination of the table that the ultracentrifuge supplies valuable information without being a precision method. Its absolute accuracy has recently been estimated as 20%.[22] For example, for egg albumin the presently accepted value for the molecular weight is 45,000 [23] $\pm$ 2000, or 45,160,[24] whereas the earlier results were 34,200 to 34,500, and the two values given in Table 16 are 43,800 and 40,500, respectively, and the latest value from light scattering is 34,500. The actual molecular weight of egg albumin, including water of hydration, is 61,400.[25] Several writers have analyzed Svedberg's hypothesis that the molecular weights are multiples of 17,600 by statistical analysis and conclude that it is invalid.[26]

Under certain conditions, such as extreme change in pH, many of the proteins break into simple fractions such as one half and one eighth, and sometimes these changes are reversible. Insulin contains a "submolecule" of molecular weight about 12,000, comprising about 106 residues, 102 of which have been analyzed. Each submolecule has 4 chains; the terminal amino groups are 2 residues each of glycine and phenylalanine.[27]

Centrifuges

Centrifuges are most useful tools in the study of colloidal systems. They are indispensable in many cases for separating emulsions or for concentrating coarser suspensions. Even the finest emulsions of colloidal dimensions can be separated by using a suitably designed parabolic insert in the directly air-driven rotor described above. Bechhold and Schlesinger [28] in a valuable paper on the uses of the centrifuge showed that if a centrifuge retaining sedimented particles on the bottom (as, for example, when it is lined with filter paper) is used and the centrifugation is carried out until most of the particles have been sedimented, the power of discrimination between particles of different sizes greatly exceeds anything possible in an ultracentrifuge. For

TABLE 16

MOLECULAR CONSTANTS OF PROTEINS

PROTEIN	s_{20}	D_{20}	M_s	M_e	M_{calc}	f/f_0	Isoel. Point	$du/dpH_0 \times 10^6$
Erythrocruorin (*Lampetra*)	1.87	10.65	17,100	19,100	$17,600 = \frac{1}{2} \times 35,200$	1.2	5.60	3.2
Lactalbumin α	1.9	10.6	17,500	...		1.2	5.12	6.7
Cytochrome C	1.89	10.13	15,600	...		1.3	9.7	
Myoglobin	2.04	11.25	17,200	17,500		1.1	7.0	7.0
Gliadin	2.00	6.72	26,000		$26,000 = 0.77 \times 35,200$	1.6		
Hordein	2.0	6.5	27,000					
Zein	1.9	4.0	35,000	33,600	$35,200 = 1 \times 35,200$			
Erythrocruorin (*Arca*)	3.46			31,400		1.0		
Erythrocruorin (*Chironomus*)	2.00			37,900		1.6	5.40	3.6
Lactoglobulin	3.12	7.27	41,800	39,200		1.2	5.19	11.9
Pepsin	3.3	9.00	35,500	35,100		1.1		
Insulin	3.47	8.20	40,900	35,000		1.1		
Bence-Jones α	3.55		37,700			1.0	5.20	5.8
Bence-Jones β	2.85	7.33				1.3	5.46	3.5
Egg albumin	3.55	7.76	43,800	40,500		1.1	4.55	10.4
CO-hemoglobin (horse)	4.5	6.3	69,000	68,000	$70,400 = 2 \times 35,200$	1.2	6.92	7.2
CO-hemoglobin (man)	4.5	6.9	63,000			1.2	7.09	6.4
Serum albumin (horse)	4.5	6.17	70,200	66,900		1.2	4.80	9.1
Yellow ferment	5.76	6.28	82,800	77,800		1.2	5.22	6.4
Serum globulin (horse)	7.1	4.05	167,000	150,000	$140,800 = 4 \times 35,200$	1.4	5.1 α, β 6.0 γ	
Phycocyan (*Ceramium*, dissoc. component)	6.2	4.58	131,000	146,000		1.4	4.85	10.2
Phycoerythrin (*Ceramium*)	12.0	4.00	290,000	292,000	$282,000 = 8 \times 35,200$	1.2	4.25	14.2
Phycocyan (*Ceramium*, main component)	11.4	4.05	272,000	273,000		1.2	4.85	10.2
Edestin	12.8	3.93	309,000	...		1.2		
Excelsin	13.3	4.26	294,000	...		1.1		
Amandin	12.5	3.62	329,000	...		1.3		

Substance								
Erythrocruorin (*Daphnia*)	16.3				422,000 = 12 × 35,200	1.1		
Hemocyanin (*Pandalus*)	17.4			397,000		1.2		
Hemocyanin (*Palinurus*)	16.4	3.4	446,000	447,000		1.5	5.05	8.1
Hemocyanin (*Helix pomatia*, dissoc. component)	12.1	2.23	503,000			1.4	4.49	10.7
Hemocyanin (*Busycon*, dissoc. component)	13.5	3.29	379,600			1.9	4.6	14
Hemocyanin (*Eledone*, dissoc. component)	10.6	2.25	440,000			1.5		
Thyroglobulin	19.2	2.65	628,000	650,000		1.2	4.58	11
Hemocyanin (*Nephrops*)	24.5	2.79	820,000		845,000 = 24 × 35,200	1.3	4.64	13.3
Hemocyanin (*Homarus*)	22.6	2.78	752,000			1.9	4.95	18
Hemocyanin (*Helix pomatia*, dissoc. comp.)	16.0	1.82	814,000	803,000		1.8	5.05	8.1
Hemocyanin (*Helix nemoralis*, dissoc. comp.)	16.6	1.92	799,000	797,000		1.4	4.63	11.4
Erythrocruorin (*Planorbis*)	33.7	1.96	1,636,000	1,539,000	1,690,000 = 48 × 35,200	1.2	4.77	10.6
Hemocyanin (*Calocaris*)	34.0			1,329,000				
Hemocyanin (*Octopus*)	49.3	1.65	2,785,000	3,000,000	2,960,000 = 84 × 35,200	1.4	4.6	14
Hemocyanin (*Eledone*)	49.1	1.64	2,790,000			1.4		
Erythrocruorin (*Arenicola*)	57.4				3,380,000 = 96 × 35,200	1.3	4.56	16
Chlorocruorin (*Spirographis*)	55.2							
Hemocyanin (*Rossia*)	56.2	1.58	3,316,000			1.4	5.28	12.6
Erythrocruorin (*Lumbricus*)	60.9	1.81	3,140,000	2,946,000		1.2		
Hemocyanin (*Helix pomatia*, main comp.)	98.9	1.38	6,630,000	2,680,000	6,760,000 = 192 × 35,200	1.2	5.05	8.1
Hemocyanin (*Busycon*, main comp.)	101.7					1.2	4.49	10.7
Hemocyanin (*Busycon*, aggregation comp.)	130.4				10,140,000 = 288 × 35,200			
Horse sickness virus	286	0.8	44,500,000			1.2	4.40	10.7

example, if in a given time a particle sediments under conditions of free convection until the residual liquid is reduced to one tenth of its original concentration, the reduction of concentration for particles double that size would be a thousand times greater.

A necessary tool in the study of viruses in order to obtain them in visible amounts is the high-speed, air-driven, angle centrifuge described by Bauer and Pickels,[29] which consists of a solid block containing diagonal holes into which the centrifuge tubes are placed. This is run at a maximum centrifugal force of 95,000 times gravity.

Hauser [30] and others have used a Sharples supercentrifuge for fractionating such materials as clay. Harvey [31] has described centrifuges with remarkably sharp microscopic observation while they are running at 12,000 times gravity.

Uses of Centrifuges

In colloid science centrifuges are used not only for separation and fractionation of emulsions and sediments, and sampling the intermicellar fluid and distinguishing between colloids and coarser suspensions or emulsions, but they are also used for purification and concentration.

REFERENCES

Ultracentrifuges

SVEDBERG, T. and PEDERSEN, K. O., *The Ultracentrifuge*, Oxford University Press, 1940; MACINNES, D. A., *et al.*, "The Ultracentrifuge," *Annals New York Acad. Sci.*, **43** (1942)

Text References

1. MASON, M. and WEAVER, W., *Phys. Rev.*, **23**, 412 (1924)
2. ARCHIBALD, W. J., *Phys. Rev.*, **53**, 746 (1938); **54**, 371 (1938)
3. ARCHIBALD, W. J., *J. Phys. Coll. Chem.*, **51**, 1204 (1947)
4. ARCHIBALD, *loc. cit.*
5. SVEDBERG, T., *Nature*, **139**, 1051 (1937); *Ind. Eng. Chem.* (Anal. Ed.), **10**, 113 (1938)
6. SVEDBERG, T., *Ber.*, **67**, 117 (1934)
7. HENRIOT and HUGUENARD, *Compt. rend.*, **180**, 1389 (1925); *J. phys. radium*, **8**, 443 (1927)
8. MACHATTIE, L. E., *Rev. Sci. Instr.*, **12**, 429 (1941)
9. MCBAIN, J. W. and O'SULLIVAN, C. M., *J. Am. Chem. Soc.*, **57**, 2631 (1935)
10. BISCOE, J., PICKELS, E. G. and WYCKOFF, R. W. G., *J. Exp. Med.*, **64**, 39 (1936)
11. MCBAIN, J. W. and LEWIS, A. H., *Trans. Farad. Soc.*, **36**, 381 (1940)
12. ELFORD, W. J., *Brit. J. Exp. Path.*, **17**, 399, 422 (1936)
    MCINTOSH, J. and SELBIE, F. R., *ibid.*, **18**, 162 (1937)
13. MCBAIN, J. W. and ALVAREZ-TOSTADO, C., *J. Am. Chem. Soc.*, **59**, 2489 (1937)
14. MCBAIN, J. W. and LEYDA, F. A., *Acta Physicochim.* (USSR), **14**, 421 (1941)

15. MCBAIN, J. W. and MCBAIN, M. E. L., *Proc. Roy. Soc.*, 139A, 26 (1933)
16. MCBAIN, J. W. and STUEWER, R. F., *Kolloid-Z.*, 74, 10 (1936)
17. MCINTOSH and SELBIE, *op. cit.*, p. 162
18. MCBAIN, J. W., *J. Am. Chem. Soc.*, 58, 315 (1936)
19. LANSING, W. D. and KRAEMER, E. O., *J. Am. Chem. Soc.*, 58, 1471 (1936)
20. ADAIR, G. S. and ADAIR, M. E., *Proc. Roy. Soc.*, 120B, 422 (1936)
21. BULL, H. B. and COOPER, J. A., Publ. #21, A.A.A.S. (1943) p. 150
22. BERNAL, J. D., *Nature*, 159, 581 (1947)
23. GUTFREUND, H., *Nature*, 153, 406 (1944)
24. BULL and COOPER, *loc. cit.*
25. BULL and COOPER, *loc. cit.*
26. NORRIS, A. D., *Nature*, 157, 408 (1946)
27. CHIBNALL, A. C., in a discussion following BERNAL, J. D., *Nature*, 159, 581 (1947)
28. BECHHOLD, H. and SCHLESINGER, M., *Biochem. Z.*, 236, 388 (1931)
29. BAUER, J. H. and PICKELS, E. G., *J. Exp. Med.*, 64, 503 (1936)
30. HAUSER, E. A. and REED, C. E., *J. Phys. Chem.*, 40, 1169 (1936); 41, 911 (1937)
31. HARVEY, E. N., *Science*, 77, 430 (1933)

SOAPS AND OTHER COLLOIDAL

ELECTROLYTES

Colloidal electrolytes are among the most numerous of colloids. In aqueous solution they exhibit the properties of both colloids and electrolytes. In defining this class, the author in 1912 [1] explained it as representing the substitution of one of the ions of an electrolyte by charged colloidal particles.

Many are association colloids; that is, the colloidal particles are formed by association of the ions, predominantly those that are surface active. Furthermore, this association is spontaneous and reversible and the colloidal particles are self-organizing and stable. The author considers it a prerequisite for their formation that a polar or soluble group of each molecule is exposed to the solvent and thus held in solution, whereas the insoluble portions of the molecules are thrust out of solution and cluster together inside the colloidal particles or micelles. Some colloidal electrolytes, such as proteins, are high polymers carrying dissociable groups which are exposed to the solvent.

The best known and most studied of the colloidal electrolytes are the soaps and synthetic detergents, followed by the dyes, which have received far less study by physical chemical methods.

The colloidal electrolytes are often classified according to the ion that gives them these distinctive properties. The soaps and many of the synthetic detergents are anion active. The high molecular weight alkylamine hydrochlorides, alkyl pyridinium chlorides or bromides, and alkylammonium salts are cation active. A very few colloidal electrolytes are both anion and cation active. Such a combination tends

to be insoluble. Indeed Lambert employs such mutual precipitation of an anionic with a cationic agent as a means of analysis for either.[2]

Then there is a large group of closely related compounds which do not form ions but do form similar colloidal particles, and cannot therefore be called colloidal electrolytes; these are known as nonionic detergents. They have the practical advantage of being indifferent to hard water and to salts, acids, and bases. These include monoglycerides and similar fatty derivatives of polyglycols with such compounds as the condensation products of alkyl phenols with polyethylene oxides. If the latter are sulfated they become anion active. Other anion active detergents are the long-known Turkey red oils (sulfonated ricinoleates), bile salts, sulfonated fatty acid esters and amides, sulfonated alcohols, polyether sulfonates, and alkyl aryl sulfonates, such as triisopropyl naphthalene sulfonate and sodium benzene sulfonates substituted by hydrocarbons from petroleum.

This gives but a faint indication of the number of such possible compounds, one patented trade name alone being said to include over a million possible chemical formulas. Many of the synthetic detergents have become very valuable for specific applications, but soap is still the cheapest and most efficient of all the known surface agents, especially for use with soft water and without the presence of incompatible chemicals such as acids, salts, and certain dyes.

A large number of further synthetic detergents become commercially available each year. In 1947 there were produced 750 million pounds of synthetic detergents, including those originally modeled on soap but avoiding hydrolysis, together with the very many subsequent ones. The earliest and still well-known representatives are the sodium alkyl sulfates and sulfonates, such as sodium dodecyl sulfate or the corresponding sulfonate. These show properties completely similar to those of the soaps, although they do not hydrolyze. This point was first demonstrated with hexadecyl sulfonic acid.[3]

Not only are colloidal electrolytes important in themselves, they have offered the advantage over most other colloids of having a definite, long-established chemical formula. Hence the colloidal particles are formed by association of these simple molecules or of their ions. Colloidal electrolytes show all the phenomena of colloids and exhibit a remarkably varied behavior which, however, has been shown to be subject to quantitative control and to be perfectly reproducible. It was upon this basis that the author was able to point out the true stability of such colloidal solutions fulfilling the same strict thermody-

namic requirements as do solutions of potassium chloride or sucrose, producing reversible equilibria within the solution and obeying the phase rule for heterogeneous equilibria.[4]

In addition, under special circumstances some unstable and metastable phenomena can be produced. Hence the colloidal electrolytes have presented unique experimental material for examination by all the quantitative methods of physical chemistry and of colloid chemistry.

The formula of sodium oleate is that of an electrolyte, NaOl, and the first important measurement to be made was that of electrical conductivity, which is the primary property of an electrolyte. However, it had to be shown that the excellent conductivity observed, which is of the same order of magnitude as that of sodium acetate, was not due to hydrolysis but must be due to the soap or colloidal electrolyte itself. Soaps exhibit a hydrolysis alkalinity which is usually less than N/1000 hydroxyl ion.[5]

The hydrolysis of soap solutions has been measured by means of indicators, the hydrogen electrode, the glass electrode, the quinhydrone electrode, ultrafiltration, conductivity, rate of catalysis, and the partition coefficient with benzene. Soap solutions made from equivalent quantities of sodium or potassium and fatty acid never produce quite enough fatty acid by hydrolysis to saturate the water, so that there is no solid precipitate or liquid emulsion of free fatty acid in the solvent.[6] Instead, the main product of hydrolysis is one or other of the acid soaps, several formulas for which are known, the most important being similar to the long known NaAc·HAc. These acid soaps are insoluble in water and appear as very fluffy turbid flocs in hydrolyzed soap solutions, but in more concentrated solutions they are solubilized by the soap and leave the soap solution completely clear and transparent. The sodium soaps are not soluble in benzene but the acid soaps are freely soluble therein.

Since colloidal electrolytes do conduct, and conduct excellently,[7] it is more fundamental to investigate the thermodynamic properties.[8] The lowering of freezing point,[9] lowering of vapor pressure, rise of boiling point, and magnitude of osmotic pressure are properties which reflect the number of particles of all sizes present in a given volume of solution, whether these particles be ions, molecules, or colloid. If association occurs it must result in a corresponding diminution in the number of ions or molecules in the solution because two or more ions or molecules are being replaced by one colloidal particle.

The most convenient method to measure the thermodynamic properties, and thus to ascertain the effective number of particles, is

to determine the osmotic coefficient. In the case of freezing point lowering in water, the osmotic coefficient g is the freezing point lowering θ divided by the freezing point lowering which would be exhibited for complete dissociation into ideal ions — in the case of NaOl the two ions, the sodium ion and the oleate ion. Therefore the osmotic coefficient g equals

$$\theta/(2m \times 1.858)$$

where m is the number of gram equivalents per 1000 grams of water, 2 is the number of ions for complete dissociation into ions, and 1.858 is the ideal lowering for a one equivalent N solution of a nonelectrolyte. The osmotic coefficient g is almost identical with Lewis and Randall's $1 - j$ or with van't Hoff's $\iota/2$. McBain and collaborators in 1911 [10] measured the electrical conductivity and osmotic coefficient of a large number of colloids. The significance of these measurements, namely, the demonstration of the existence of self-associating colloidal particles exhibiting electrical conductivity, was pointed out. Since then very numerous measurements have been supplied by many laboratories throughout the world.

FIGURE 17·1 Equivalent conductivity at 25° of undecyl, dodecyl, and tetradecyl sulfonic acid over a wide concentration range. Taken from McBain, Dye, and Johnston, Text Ref. No. 11.
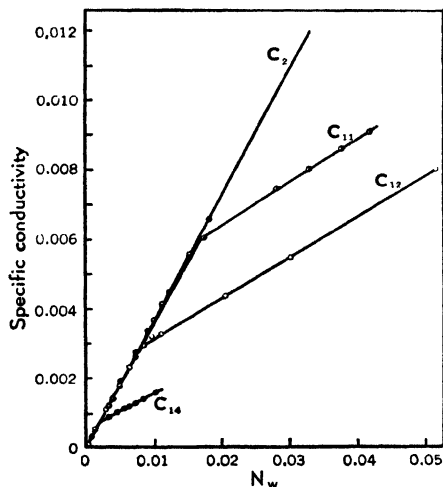
FIGURE 17·2 Deviation of specific conductivity from that for a fully dissociated electrolyte. Taken from McBain, Dye, and Johnston, Text Ref. No. 11.

Figure 17·1 presents the data of M. E. L. McBain, Dye, and Johnston [11] for the equivalent conductivity at 25° C of undecyl, dodecyl, and tetradecyl sulfonic acid over a wide concentration range. Figure 17·2 presents the same data in different form. Here the specific conductivity is plotted against the concentration, and it is shown that the conductivity in all three cases merges in dilute solution with the

conductivity for the ordinary electrolyte, ethyl sulfonic acid, while departing from it at higher concentrations. This method of plotting the data was first employed by Wright, Abbott, Sivertz, and Tartar.[12]
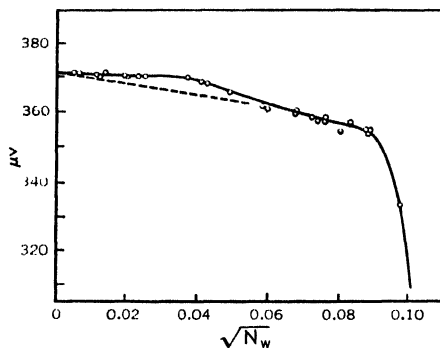
FIGURE 17-3 Equivalent conductivity of very dilute solutions of dodecyl sulfonic acid. Taken from McBain, Dye, and Johnston, Text Ref. No. 11.

Figure 17.3 shows an enlarged graph of the data of McBain, Dye, and Johnston for equivalent conductivity of lauryl sulfonic acid in the very dilute range of solution. The dotted straight line is the Onsager slope,

$$\mu_\sigma = \mu_\infty - [0.2774\mu_\infty + 59.78]\sqrt{N_r}$$

The main feature of the electrical conductivity is the approximate agreement with that to be expected for a fully dissociated electrolyte, as indicated in Figures 17·2 and 17·3 — but only in the most dilute solutions. When a certain concentration is reached (0.008 N for dodecyl sulfonic acid), the so-called critical concentration (the critical concentration cannot be sharp, according to the principle of mass action, but must extend over a narrow range of concentration), the equivalent conductivity falls rapidly with increasing concentration and subsequently passes through a minimum and rises again. The concentration for the completion of formation of micelles is ten or more times greater than the critical concentration for their initiation.

Figure 17.4 gives the osmotic coefficient of dodecyl sulfonic acid. The straight line in the figure is the Debye-Hückel-Onsager slope for an ideal, fully dissociated electrolyte. The osmotic coefficient likewise shows an approximation of the behavior of an ideal, fully dissociated electrolyte in very dilute solution, but again, when reaching a certain concentration, the osmotic coefficient falls off sharply, proving the disappearance of ions to form colloidal particles or micelles. Since the

osmotic coefficient falls so low, these particles must be of colloidal dimensions. Indeed, as was pointed out nearly forty years ago by the author, there is not enough osmotic activity to account for the observed conductivity even if the whole of the osmotic activity be attributed to sodium ions. A large fraction of the conductivity remains to be explained only by conducting colloidal particles. Hence, all investigators in this field now are agreed that the outstanding feature of the colloidal electrolytes is the formation, above a certain concentration, of conducting colloidal particles. Their nature and the still existing uncertainty as to their size and number will be discussed later. We
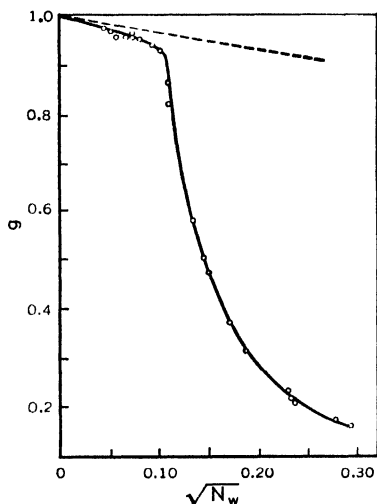
FIGURE 17·4 The osmotic coefficient of dodecyl sulfonic acid. The straight line is the Debye-Hückel-Onsager slope for ideal fully dissociated electrolyte. Taken from McBain, Dye, and Johnston, Text Ref. No. 11.

then can comment upon the fact that the conductivity of lauryl sulfonic acid in very dilute solution rises well above that for a fully dissociated electrolyte, and that correspondingly in Figure 17·4 the osmotic coefficient in dilute solution still lies slightly below the value for complete ionization.

Lottermoser and Püschel [13] were the first to present extensive data on nonhydrolyzable synthetic detergents which emphasize this behavior in dilute solutions. Some of their conductivity results are reproduced in Figures 17·5 and 17·6. They show the comparatively minor change of critical concentration with temperature and its great dependence upon the molecular weight of the colloidal electrolyte.
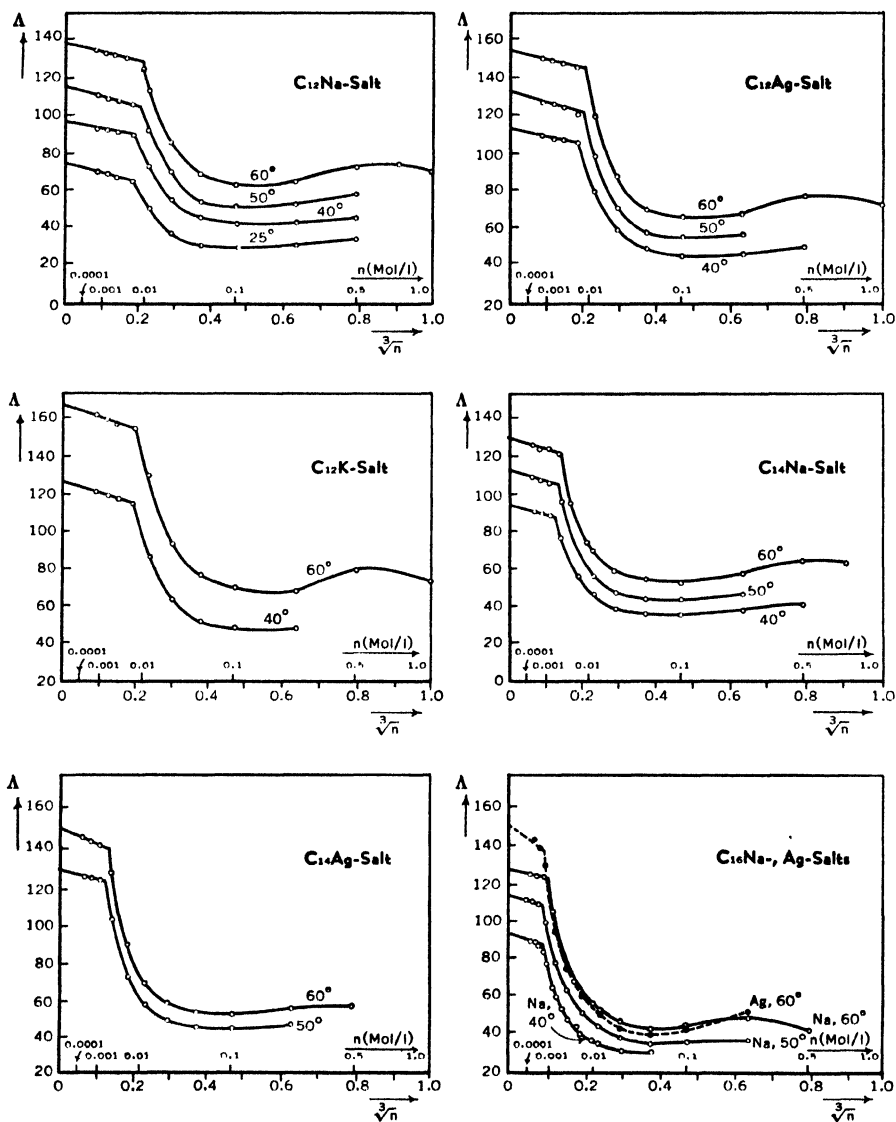
FIGURE 17·5 Some conductivity data of Lottermoser and Püschel, Text Ref. No. 13.

Many similar conductance curves have since been obtained by Tartar and his collaborators,[14] by Ralston and Hoerr,[15] and many others,[16] using alkyl ammonium salts and other kinds of detergents.

Figure 17·7 shows the gradation in shape of the family of conductivity curves, obtained by the author's collaborators,[17] for a single
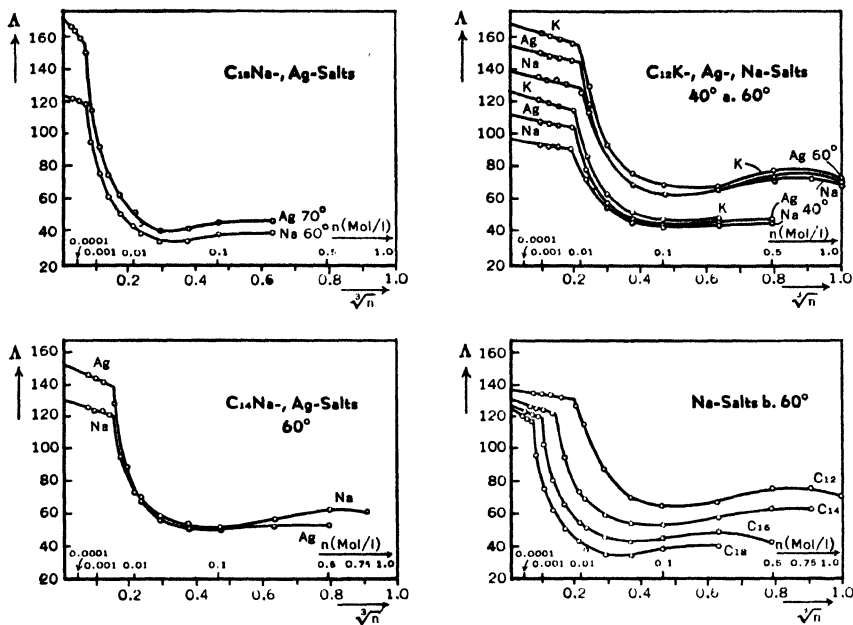
FIGURE 17-6 Some conductivity data of Lottermoser and Püschel, Text Ref. No. 13.

homologous series of saturated fatty acids. In this study no particular attention was paid to dilute solutions because of the uncertainty introduced there by hydrolysis. This hydrolysis can be completely suppressed, however, by the addition of a very slight amount of free alkali. The figure includes also some results for potassium oleate and sodium oleate.

Gonick [18] points out that the conductivity multiplied by the square root of the concentration, when plotted against the concentration for any colloidal electrolyte, yields a straight line,

$$\Lambda \sqrt{C} = A + BC$$

where A and B are constants. The intercept A depends mainly upon

FIGURE 17-7 Equivalent conductivity in mhos for potassium salts of the saturated fatty acids, from the acetate, Ac, to the stearate, Str, together with measurements of oleates, Ol. Taken from Bunbury and Martin, Text Ref. No. 7.

g
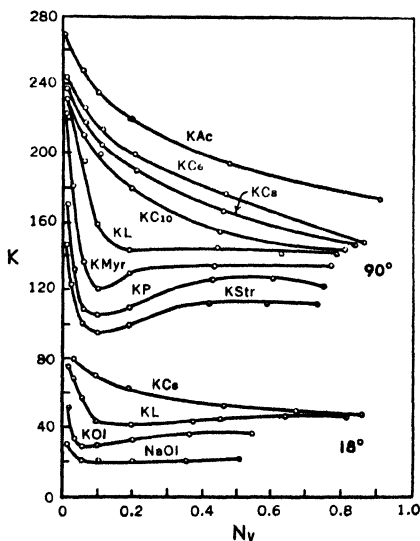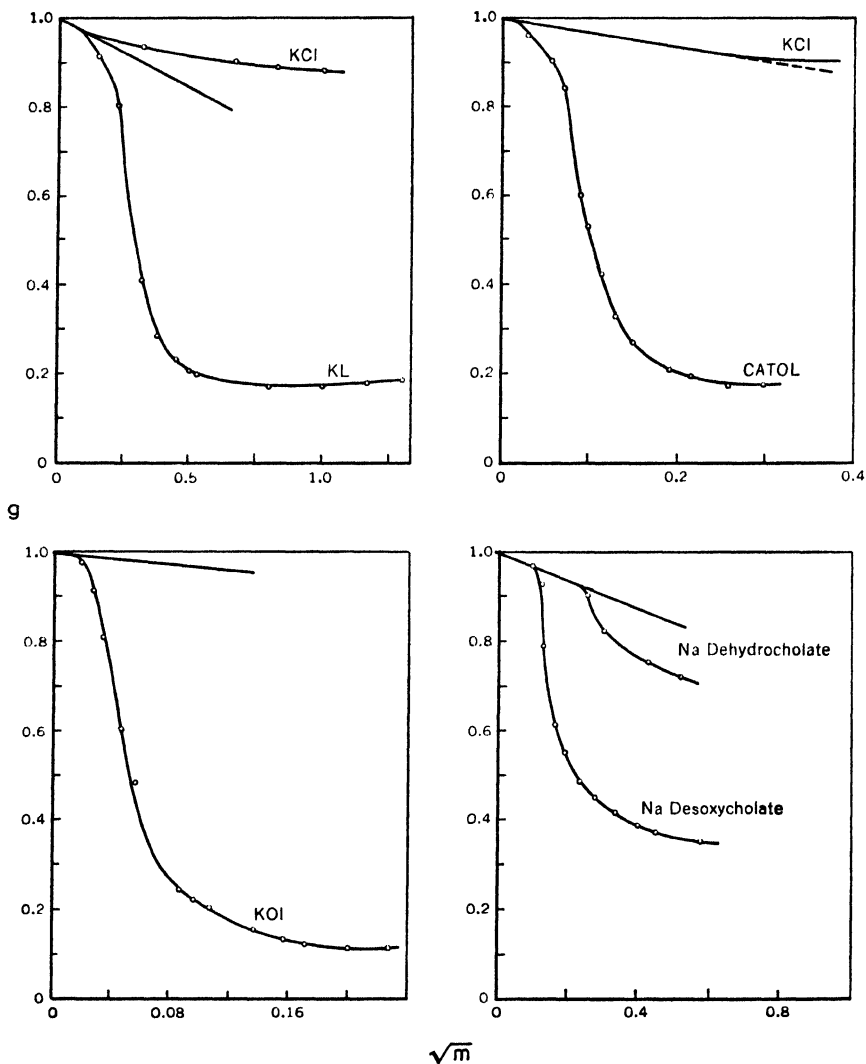
FIGURE 17-8 Osmotic coefficients of various colloidal electrolytes as indicated. The straight line in each case is the Debye-Hückel-Onsager slope. Taken from work done by Bolduan, by Johnston, and by Brady in the Stanford laboratories.

the anion, and the slope B at a given temperature depends only on the cation. The equivalent conductivity in the region of the conductivity minimum is the sum of $A/\sqrt{C} + B/\sqrt{C}$ and it occurs at a concentration equal to A/B.

In contrast to studies of electrical conductivity, which have come from many laboratories, the thermodynamic data for osmotic coeffi-

cients have been provided chiefly, although not entirely, by the author's collaborators beginning with the measurements of the dew point lowering by McBain and Salmon,[19] very numerous freezing point lowerings, and lowerings of vapor pressure by several methods. Some typical data are given in Figure 17·8. Comparison is shown between the behavior of these colloidal electrolytes and that of the typical electrolyte potassium chloride and the theoretical Debye-Hückel-Onsager slope.
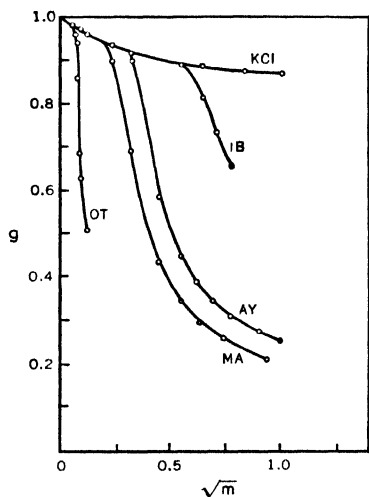
FIGURE 17·9 Osmotic coefficients of a series of aerosols of different molecular weight, from IB (diisobutyl) to OT (dioctyl). Taken from McBain and Bolduan, *J. Phys. Chem.*, **47**, 94 (1943).

FIGURE 17·10 Osmotic coefficients of all colloidal electrolytes brought to coincidence on one or other of three curves depending on whether they are straight chain, branched chain, or polycyclic, by dividing the concentrations for each by $m_g = 0.5$, the molality at which $g = 0.5$. The dashed line represents the limiting slope at which the solution would fall into separate layers. Taken from McBain and Brady, Text Ref. No. 20.

Figure 17·9 gives the osmotic coefficients of a series of aerosols, that is, sodium sulfonates of diesters of succinic acid of different molecular weight from IB (diisobutyl) through AY (diamyl) and MA (dihexyl) to OT (dioctyl).

Now there is a most remarkable fact, discovered by Brady,[20] that all the osmotic data hitherto measured for colloidal electrolytes could be placed upon one of three curves, as shown in Figure 17·10, by the simple device of dividing the concentration for each by the molality at which $g = 0.5$. There is one curve each for straight chain, branched chain, and cyclic compounds.

Comparing now the data for conductivity and those for osmotic coefficient, we may look at Figures 17·11, 17·12, and 17·13 for potas-

FIGURE 17·11 Comparison of osmotic coefficient and conductivity for aqueous solutions of potassium oleate. The abscissas are the square root of concentration; the ordinates are α and g. Taken from Johnston and McBain, *Proc. Roy. Soc.*, 181A, 119 (1942).
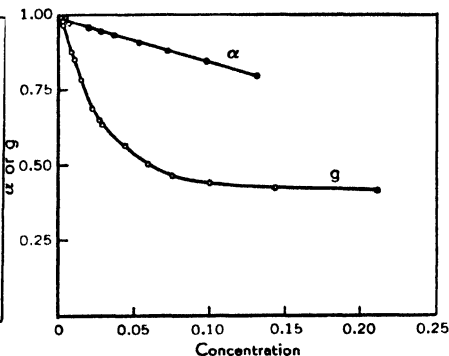
FIGURE 17·12 Comparison of osmotic coefficient and conductivity for sodium desoxycholate, at 0°. Taken from McBain and Johnston, *Proc. Roy. Soc.*, 181A, 129 (1942).

FIGURE 17·13 The osmotic coefficient and the "Arrhenius coefficient," $\alpha = \lambda/\lambda_0$, of aqueous hexanolamine oleate plotted against molality; those for potassium chloride are shown for comparison. Taken from Gonick and McBain, *J. Coll. Sci.*, 1, 127 (1946).
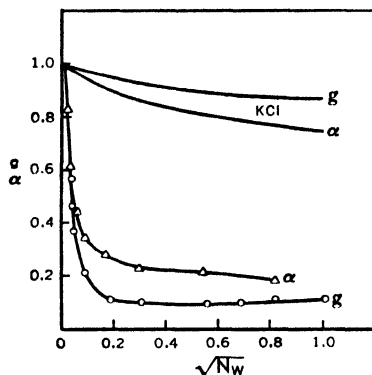
sium oleate, sodium desoxycholate, and hexanolamine oleate, respectively. It is typical that the conductivity ratio is not as greatly lowered as the osmotic coefficient, the reverse of what is observed for potassium chloride, as is shown in Figure 17·13. The conductivity shows little or no sign of formation of colloidal electrolyte in such cases as a bile salt, or an aerosol, or the dye metabenzopurpurin, investigated by Hartley,[21] where other measurements, such as transport number, show the presence of micelles.

Brady noted that if all the freezing point lowering for lauryl sulfonic acid, Catol 607, or potassium laurate is ascribed to the simple ion, namely, hydrogen ion, chloride ion, and potassium ion, respec-
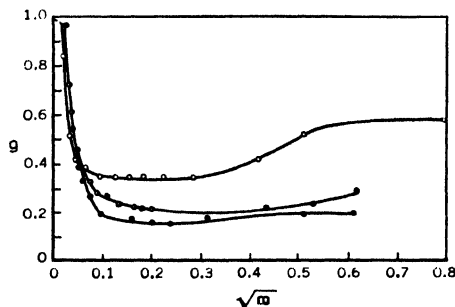
FIGURE 17·14 Osmotic coefficients of some nonionic detergents; $\bigcirc$ = nonethyleneglycol monolaurate; $\oplus$ = Detergent "X"; $\bullet$ = Triton X-100. Taken from Gonick and McBain, *J. Am. Chem. Soc.*, **69**, 334 (1947).

tively, the observed conductivity is 24%, 29%, and 30% respectively greater than that so calculated, thus again emphasizing the fact that the colloidal micelles carry a substantial fraction of the total current.

It is interesting to note that the osmotic coefficient of nonionic detergents exhibits completely similar behavior to that of colloidal electrolytes, although of course conductivity is absent. This is illustrated in Figure 17·14.

Table 17 assembles some of the numerous data for osmotic coefficients as obtained in the author's laboratory.

Critical concentrations for the initiation of formation of micelles have been determined by the conductivity methods illustrated in Figures 17·1 and 17·2, by an analytical method using osmotic coefficients, solubilization of dyes,[22] color change of dyes in detergent solutions (Harkins and collaborators), change in partial molar volume (Bury), and refractive index (Klevens). The results of the different methods do not agree exactly, as would be expected from the fact that solubilization anticipates the concentration by promoting micelle formation, nor do they represent exact concentrations but rather a zone, as required by the mass law. Some typical results of Corrin, Klevens, and Harkins are cited in Table 18. Klevens [23] has deduced "for straight chain saturated soaps and detergents that the critical concentration depends solely on the length of the surface-active ion. Thus a C_{13} fatty acid, a C_{12} sulfonate, a C_{11} sulfonate, and a C_{12} aminehydrochloride belong to the same group."

TABLE 17 *

OSMOTIC COEFFICIENTS OF SOAPS AND DETERGENTS

K caprate, C_{10}	m	0.040	0.100	0.200	0.300	0.5	0.935
	g	0.954	0.943	0.770	0.548	0.350	0.253
Na laurate, C_{12}	m	0.043		0.081			
	g	0.593		0.374			
K laurate	m	0.020	0.051	0.102	0.152	0.263	0.651
	g	0.906	0.788	0.402	0.281	0.199	0.161
K myristate, C_{14}	m	0.012	0.021	0.062	0.197		
	g	0.66	0.453	0.243	0.127		
Na oleate	m	0.011	0.042	0.095			
	g	0.354	0.244	0.241			
K oleate	m	0.0004	0.0012	0.002	0.007	0.024	0.161
	g	0.981	0.805	0.593	0.269	0.123	0.102
K butyrate	m	0.091	0.4	0.5	0.7		
	g	0.929	1.01	1.02	1.06		
Na linoleate	m	0.062	0.116	0.340			
	g	0.143	0.132	0.111			
Na linolenate	m	0.047	0.128	0.500			
	g	0.134	0.108	0.096			
Na ricinoleate	m	0.060	0.152	0.302			
	g	0.331	0.211	0.172			
Na abietate	m	0.010	0.024	0.048			
	g	0.811	0.615	0.597			
Na decyl sulfonate	m	0.002	0.005	0.011	0.018		
	g	0.987	0.975	0.958	0.899		
Na decyl sulfate	m	0.002	0.006	0.014	0.037	0.054	
	g	1.0	0.98	0.99	0.89	0.73	
Na dodecyl sulfate	m	0.001	0.003	0.005	0.006		
	g	0.987	0.967	0.942	0.882		
Na deoxycholate	m	0.003	0.010	0.013	0.022	0.074	0.323
	g	1.00	0.92	0.84	0.68	0.45	0.34
Na dehydrocholate	m	0.020	0.050	0.100	0.281		
	g	0.980	0.903	0.813	0.718		
Benzyl trimethyl ammonium laurate	m	0.018	0.042	0.087	0.181	0.389	
	g	0.99	0.556	0.371	0.300	0.288	
Morpholine laurate	m	0.042	0.087	0.180	0.379	0.592	
	g	0.521	0.395	0.294	0.215	0.203	
Monobutyl biphenyl sodium monosulfate	m	0.034	0.050	0.103	0.178	0.368	0.566
	g	0.779	0.784	0.713	0.687	0.628	0.600
Monobutyl phenylphenol sodium monosulfonate	m	0.025	0.050	0.105	0.633	0.568	
	g	0.940	0.840	0.745	0.268	0.580	
Lauryl pyridinium chloride	m	0.050	0.100	0.200			
	g	0.355	0.311	0.287			
Cetyl pyridinium chloride	m	0.005	0.010	0.042			
	g	0.685	0.412	0.120			

TABLE 17 *(continued)* 253

Lauryl benzyl dimethyl ammonium chloride	m	0.032	0.064	0.168	0.358		
	g	0.28	0.204	0.163	0.150		
Cetyl dimethyl benzyl ammonium chloride	m	0.005	0.05	0.10	0.20		
	g	0.62	0.101	0.105	0.107		
Myristyl trimethyl ammonium chloride	m	0.050	0.100	0.200			
	g	0.181	0.152	0.137			
Cetyl trimethyl ammonium bromide	m	0.005	0.05	0.10	0.20		
	g	0.91	0.057	0.089	0.076		
Aerosol OT	m	0.0036	0.0046	0.0066	0.0076	0.011	
	gm	0.958	0.864	0.692	0.629	0.509	
Aerosol MA	m	0.050	0.100	0.200	0.400	0.555	0.890
	g	0.902	0.697	0.438	0.300	0.262	0.211

* Values from Johnston and McBain, *Proc. Roy. Soc.*, 1942, **181**, 119; McBain and Bolduan, *J. Phys. Chem.*, 1943, **47**, 94; Fineman and McBain, *J. Phys. Coll. Chem.*, 1948, **52**, 881; Cushman, Brady and McBain, *J. Coll. Sci.*, 1948, **3**, 425.

In alcohols,[24] soaps dissolve as simple, partially dissociated electrolytes, not colloidal electrolytes. The transition from aqueous to nonaqueous solutions has been followed by several investigators.[25]

Two significant properties of colloidal electrolytes are their rates of diffusion, integral and differential (diffusion at a definite concentration), and electrolytic transport number or Hittorf number. The latter has often been misinterpreted through a confusion between electrical and material transport number.

The total current passing through a solution is divided amongst all ions and other carriers present, in proportion to their intrinsic mobilities and their respective concentrations, as was described in

TABLE 18

CRITICAL CONCENTRATIONS FOR THE INITIAL FORMATION OF MICELLES
IN SOLUTIONS OF SOAPS AND DETERGENTS

Detergent	Temperature	Critical Concentration
Na octyl sulfonate	25°	0.15 m
Dodecyl ammonium chloride	room	0.01 or 0.06
Na decyl sulfonate	25°	0.04
K dehydroabietate	26°	0.03
K laurate	26°	0.02
Decyl trimethyl ammonium bromide	room	0.06 or 0.01
Na dodecyl sulfonate	34°	0.01
K myristate	26°	0.006
Na dodecyl sulfate	26°	0.006
Na tetradecyl sulfate	43°	0.003
K oleate	26°	0.001
Na cetyl sulfonate	50°	0.0008
Na cetyl sulfate	36°	0.0004
K dilinoleate	26°	0.0002

Chapter 13. Hence, the *electrical* transport number of any species, whether ionic or colloidal, must be a *positive fraction* less than unity, and all must add up to unity. On the other hand, if there are m chemical equivalents of any kind of matter associated with one electrical charge, then the *material* transport will be m times the *electrical* transport. Hence the material transport number or Hittorf transference value may greatly exceed unity for a particular colloidal species. Thus a high material transport does not mean a high conductivity but often the reverse. Suppose, as a purely hypothetical example, in a particular concentration of potassium laurate there might be present only a colloidal particle of empirical composition $K_{24}L_{27}^{-3}$ and 3 simple K^+ ions. In this composition there are 9 equivalents of laurate to one electrical charge. It has been shown by Gonick [26] that the *electrical* transport number of the micelle would be $+ 0.227$ and the electrical transport number of the potassium ion would be $+ 0.773$. However, it is the *material* transport number of laurate ions that is actually measured, and this would be 9 times $+ 0.227$ or $+ 2.04$, which appears as the material transport number of laurate radical, that is, the number of gram equivalents of laurate radical carried through any cross section toward the anode per Faraday of current. Hence, although the electrical transport number of the simple potassium ions is $1 - 0.227 = + 0.773$, there is more potassium being carried toward the anode in the colloid than potassium ions migrating to the cathode, and the *material* transport number of the potassium appears then as $- 1.04$. This is made up of $+ 0.773$ potassium atoms carried in the simple ions toward the cathode, minus the 8 times 0.227 potassium atoms or -1.816 carried in the opposite direction in the micelle, the difference being -1.043 *material* transport of potassium ion. It is seen that the large transport of laurate toward the anode per Faraday of current does not mean a correspondingly high conductivity because there are only 3 charges on a micelle with 27 laurate radicals and the conductivity is cut down to about one ninth of that for complete ionization at infinite dilution.

Similarly the osmotic activity, instead of corresponding to potassium ions plus laurate ions for complete dissociation, is reduced to the 3 ions and 1 micelle, and the osmotic coefficient will therefore be about 0.07.

Actual values for the dependence on concentration of transport number, differential diffusion coefficient and integral diffusion coefficients of solutions of dodecyl sulfonic acid as determined by M. E. Laing McBain,[27] are given in Figure 17·15.

It is significant and important that the transport number passes through a maximum in N/20 solution. It is evident that at this concentration a change in the kind of average micelle present takes place. Hence more than one kind of micelle must exist. The differential diffusion coefficient also passes through a minimum at the same concentration, but at this concentration the conductivity is continuing to fall.

FIGURE 17·15 Electrolytic migration, or Hittorf or transport number, of lauryl sulfonate radical in solutions of dodecyl sulfonic acid compared with the diffusion data. Taken from McBain, Text Ref. No. 27.

T.N =Transport number
I.D. = Integral diffusion
D.D =Differential diffusion
C.C = Critical concentration
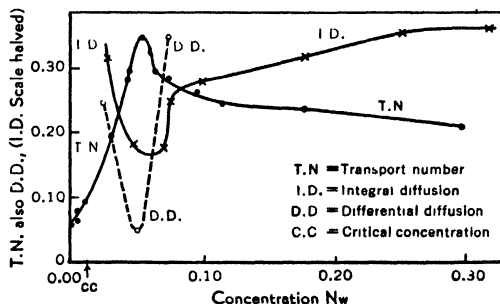
Figure 17·16 gives the transport data for sodium oleate,[28] which may be compared with the conductivity data in Figure 17·7 and the osmotic data in Figure 17·11. Figure 17·17 gives data for cetyl pyridinium and cetyl trimethyl ammonium bromides. It should be noted from the previous discussion that the equivalent *conductivities* of the component radicals are all *positive* and cannot be calculated from the transport numbers and total conductivities without a knowledge of the formulas of the carriers present.

The ultrafiltration data of McBain and Jenkins [29] likewise show that there must be different kinds of micelles simultaneously present.

The author has long emphasized that a distinctive feature of colloidal electrolytes is that the ionic strength principle does not apply to their mixtures with other electrolytes and nonelectrolytes [30] unless the colloidal electrolytes are treated as uni-univalent electrolytes.

  ' Although ionic micelles are highly charged colloidal particles their charges are superficial and spaced so far apart that they are effectively independent. In this respect, they are sharply differentiated from ordinary polyvalent ions in which the charges are coincident.[31]

The Kinds of Micelles Present in Solutions of Colloidal Electrolytes

The author was disinclined to postulate the existence of more than one kind of colloidal particle in any one solution of a colloidal electrolyte but was soon forced to recognize that such data as those referred to in connection with ultrafiltration and transport number cannot be
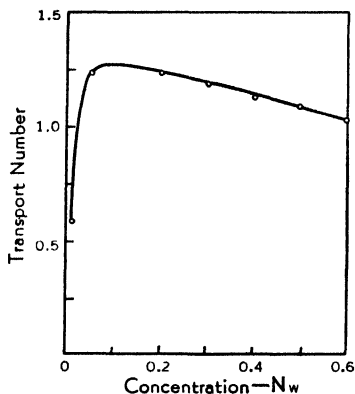
FIGURE 17·16 The transport number of solutions of sodium oleate. Plotted from data given in Laing, Text Ref. No. 28.
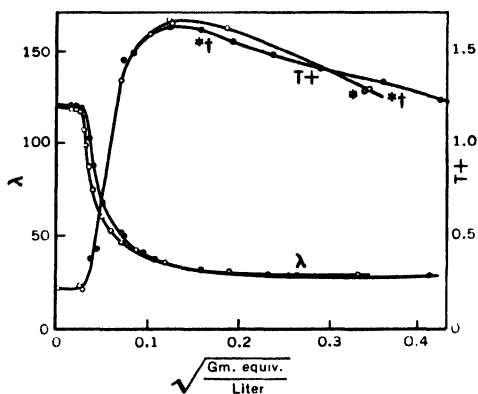
FIGURE 17·17 Total equivalent conductivity (λ) and paraffin chain transport number (T+) for cetylpyridinium (O) and cetyl trimethyl-ammonium (●) bromides at 35°C. * and † distinguish points obtained with new apparatus and BaCl₂, respectively. Taken from Hartley, Collie, and Samis, Text Ref. No. 21.

explained without assuming the existence of different kinds, sizes, shapes, and compositions of micelles. He postulated a minimum of two.

One was the so-called "nearly neutral" or slightly ionized lamellar micelle consisting of a double leaflet of soap molecules placed end to end and side by side, as is shown in Figure 1·1. That figure was prepared for the author's discourse at the Royal Institution of Great Britain on March 20, 1925, from which the following description is quoted:

> Each particle is like a pair of military hair brushes in which the bristles represent the hydrocarbon chains of the molecules arranged parallel to each other in sheets, two such layers being put together hydrocarbon to hydrocarbon. The two backs of the brushes on the outside represent the hydrated layer and the un-ionized electric double layer. A general survey of the facts with regard to the electric double layer has shown that only a minute fraction of such surfaces can ionize — hence the name "neutral" micelle.

(Compare the similar description in *Colloid Symposium Monograph*, 1926, Vol. IV, p. 8.) This kind of micelle was adopted by Harkins and collaborators in 1947.[32] Formation of this sort of micelle would account for the falling off in conductivity above the critical concentration.

The second extreme possibility of the McBain spherical ionic micelle (see Figure 17·18) would be the aggregation of a number of ions of like sign. As pointed out in 1913 by McBain,[33] such an aggregate would retain the same electrical driving force per ion per charge in an

electric field as if each ion were wholly independent, but the resistance to motion would be cut down in accordance with Stokes' law. Hence the rate of movement and the corresponding conductivity would be *increased*. Howell and Robinson [34] state:

> If n univalent spherical ions each of radius r coalesce to form a spherical micelle, carrying n charges, then it follows from Stokes's Law that the conductivity of the micelle would be $n^{\frac{1}{3}}$ times as great as that of the n individual ions.

An analogy is the rapid rate of fall of a large drop of rain in a thunderstorm as compared with the same weight of water divided into fine mist.

Enhanced conductivity above that for the same ions at infinite dilution will often be a property of small micelles, even if they consist of ions of opposite sign, if they retain a sufficiently high charge. The volume of such a particle is approximately the sum of the large surface active ions therein, and hence the resistance to movement is proportional to the radius of the aggregate or proportional to the cube root of n as compared with that of a single surface active ion. The free charge on the particle will be $n - m$ where m is the number of oppositely charged ions within the particle. It depends upon the value of $n - m$ and the conductivity of the m ions (whose conductivity is lost when they are within the micelle) whether the net effect of formation of a particular micelle results in increase or decrease of conductivity or leaves it approximately unchanged. If the micelle is heavy and not strongly charged, the conductivity will always be lessened.

FIGURE 17·18 The McBain spherical ionic micelle.

Ralston, Eggenberger, and DuBrow [35] have shown that for quaternary ammonium chlorides of sufficiently high molecular weight the conductivity at first rises above that for infinite dilution and then falls in the usual manner upon further concentration. This is like the behavior of methylene blue as found by Hartley, Collie, and Samis.[36] A similar result has been obtained for hexadecyl pyridinium iodate and still more for octadecyl pyridinium iodate.[37] The micelles responsible for this enhanced conductivity must be different from those causing the lowered conductivity in more concentrated solutions.

A few small micelles suggested by the author, including his ionic micelle, are pictured in Figure 17·18. The ionic micelle is represented

in cross section, and it is suggested that it could not contain more than about 10 fatty ions if it exists. Reychler in 1913 was the first to suggest such a spherical aggregate and it was later suggested again by Adam.[38] Stauff [39] has discussed the small micelles shown in Figure 17·19. He shares with the author the belief that every conceivable kind of aggregate that reduces the free energy (as for example by reducing hydrocarbon-water interface, promoting hydrogen bonds, etc.) will occur in solutions of colloidal electrolytes statistically in proportion to its effectiveness in reducing the free energies and in accordance with the principle of mass action. Hence those preponderantly present will be those in which this is best effected. Meyer and van der Wyk [40] pointed out that once a micelle has reached a certain size the addition of each successive ion or ion pair will be attended by approximately the same energy. Hence, they predicted a Maxwellian distribution of size of any particular kind of micelle in the same solution.

X-ray examination has given the only direct evidence so far obtained for the existence of any particular kind of micelle. Hess and Gundermann [41] explained the streaming double refraction of soap solutions by the presence of lamellar McBain micelles discussed above as a "nearly neutral" micelle. Hess and collaborators [42] have examined the x-ray diffraction of a large number of solutions of soap and syn-
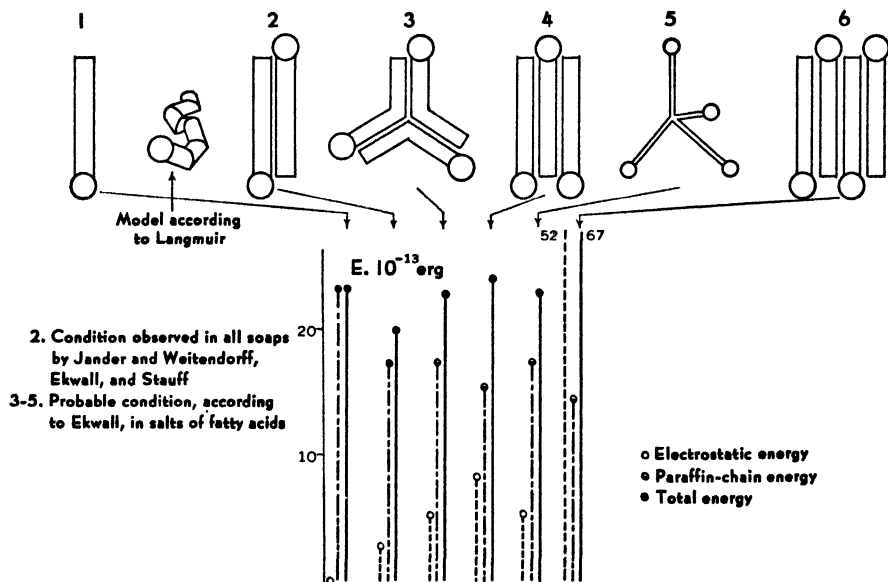
FIGURE 17·19 Stauff's suggestions for small micelles and the free energy of their formation. Taken from Stauff, Text Ref. No. 39.
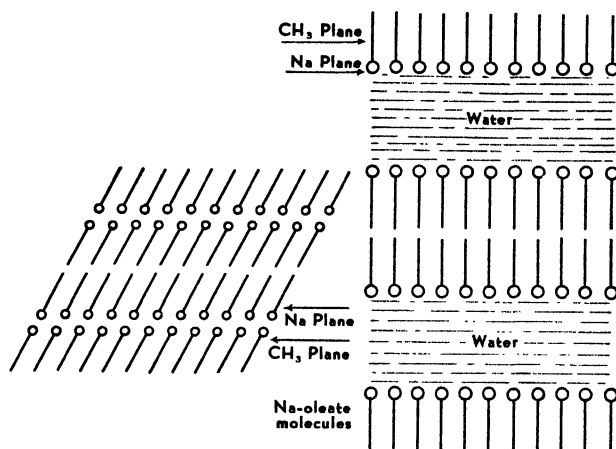
FIGURE 17·20 Comparison of the German concept of the Hess micelles with the known x-ray structure of the crystalline soap. Taken from Hess, *Fette und Seifen*, 49, 81 (1942).

thetic detergents. They find a long Bragg spacing which is more than twice the length of a soap molecule and a short spacing of 4.4–4.6 Å. This proves that the soap micelles in ordinary transparent soap solutions constitute a space lattice or x-ray diffraction grating. Figure 17·20 compares their interpretation of these large micelles with the known x-ray structure of the anhydrous crystalline soap. The molecules are laid end to end, side by side, as in the McBain lamellar micelle, but these lamellae repeat at a regular distance from each other separated by a layer of water whose thickness depends upon the concentration. It has been shown in the writer's laboratory by O. A. Hoffman and others that the molecules in the micelle are often tilted at an angle of about 55° and that they correspond to a liquid crystalline state.

Hydration of the soap micelles of this order of magnitude had previously been determined by means of ultrafiltration.[43]

Stauff [44] confirms the Hess micelles but points out that they are not crystalline and that although the molecules are close packed, side by side, this packing is irregular, as in a liquid or in a liquid crystal — that is, all the molecules do not have the zigzags of their hydrocarbon chains oriented exactly parallel as would be the case in a crystal.

A portion of one of the repeating Hess micelles with lamellar structure is given in perspective for 0.77 N potassium laurate solution in Figure 17·21. These micelles are not the only ones present because the x-ray pattern fades out at a concentration which is sometimes called a second critical concentration [45] of 3% or more. The author

believes that fragments of these micelles and small micelles must more and more take their place with increasing dilution until all have fully dissociated into ions.

In nonaqueous solution the McBain lamellar micelle, if it exists, must be turned inside out with the hydrocarbon chain out as pictured by Lawrence.[46] Association in non-ionizing solvents has been shown to occur with both ionic and nonionic detergents.[47]

S. S. Marsden, Jr., in the Stanford laboratory, has obtained x-ray evidence for such an inverted Hess micelle in solutions of triethanolamine laurate in dry benzene.

McBain's original formulations were set up before the introduction of the now accepted Debye-Hückel theory of electrolytes, but the essential conclusions as outlined in this chapter remain. Hartley [48] did important service in showing how the Debye-Hückel theory should be taken into account. From a portion only of the kinds of evidence referred to in this chapter he attempted to account for the properties of colloidal electrolytes by the use of one invariant spherical micelle, which is pictured in Figure 17·22, and several authors still use his concept. Figure 17·22 is open to criticism. It is supposed to represent a sphere of 50 liquid, close-packed, paraffin chains with their polar heads outward. The surface of such a sphere, as Ward also points out, must be largely hydrocarbon in contact with water because the 50 polar heads can cover only one quarter of the surface of such a volume. Conversely, if the heads were close-packed there would be room for only a fraction of the hydrocarbon chains within them. As Hartley points out, the chains cannot be arranged radially although their length has been taken as the radius of the sphere because the density throughout must remain that of a hydrocarbon, and the axes of the hydrocarbon chains cannot approach each other within about 5 Å since this is the distance of closest approach between two molecules. War-

FIGURE 17·21 A portion of one of the repeating Hess micelles with lamellar structure, given in perspective. Taken from Philippoff, Text Ref. No. 42.

ren,[49] using the data of Stewart, shows that the distance between adjacent hydrocarbon chains in liquid hydrocarbons is 5.0 Å and that the chains, at least up to C_{10}, do not bond appreciably. Each paraffin chain molecule is surrounded by six others parallel with it. Hence it is difficult to see how a compact sphere such as that of Hartley can be put together. Hirshfelder models, otherwise so useful, do not take sufficiently into account the surrounding van der Waals' space. This consideration has to be borne in mind in discussing high polymers.

Hartley supposes that only 10–25 of the oppositely charged ions are fully dissociated as gegenions (free ions compensating the charges on the particle) in the surrounding ionic atmosphere. For cetyl trimethyl ammonium bromide he takes the number of such charges to be 13. The soap molecules in Figure 17·18 are drawn more realistically as to scale and packing; they would owe their approximately spherical form to the mutual repulsion of uncompensated like charges.

Walton, Hiebert, and Sholtes [50] from studies of e.m.f. and freezing point depression of dodecyl benzyldimethylammonium chloride and sulfamate conclude that in the concentration range where the micelles are fully established about 60% of the chloride ions are combined in the average micelle. Like Hartley and all other investigators they find that the concentration of the simple surface active ion falls with increase of total concentration.

Klevens,[51] and likewise Harkins,[52] prefers an elongated ellipsoidal or cylindrical model as at least one form of micelle. A cylinder would be more effective packing than a sphere, but neither could compare in efficiency with lamellar packing, which more nearly approximates the known structure of a hydrocarbon liquid and exposes the minimum of hydrocarbon surface to water. Such a lamellar micelle could be so narrow as to resemble a lath or square rod. Such particles would arrange themselves parallel in hexagonal array as is evidenced by the hexagonal x-ray spot photographs obtained by Marsden [53] and Hoffman in the author's laboratory at Stanford University.
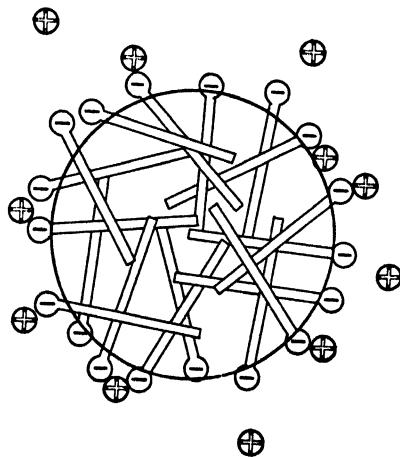
FIGURE 17·22 The Hartley spherical micelle. Taken from Hartley, *Aqueous Solutions of Paraffin Chain Salts*, p. 43, Text Ref. No. 48.

The Phases of Soaps

Soaps are remarkable in that their aqueous solutions, depending upon the temperature and concentration, exist as different phases, some being ordinary isotropic, clear, transparent, fluid solutions, colloidal except in extreme dilution; and others being liquid crystalline, anisotropic, plastic, and translucent. Among the latter are middle soap, soap boiler's neat soap, superneat, superwaxy, subneat, and neat soap phases, each of which exists over a range of concentration and temperature and is as immiscible with any other phase as would be oil and water. All these immiscible phases consist of colloidal electrolytes, as has been shown by measurements of conductivity and osmotic coefficient. The phases obey the Gibbs' phase rule. A typical phase diagram is given in Figure 17·23 for sodium stearate.[54] Similar behavior has been found for all the other alkali, amine, and hydrogen soaps studied.

X-ray examination of the more concentrated phases shows that some contain lamellar micelles of the Hess type, whose long spacings diminish with concentration; there are others whose long spacings are independent of concentration; and finally, in a few instances, of which the only completely authenticated one is dodecyl sulfonic acid, middle phase, the micelles are oriented in a hexagonal pattern of rods or narrow fibers or ellipsoids arranged parallel in such perfect hexagonal order that they can give a point diagram,[55] as contrasted with the usual powder photograph for oriented matter.

Some Other Colloidal Electrolytes

Considerations of space forbid more than a brief reference to studies of colloidal electrolytes other than the enormous family of soaps and detergents. They include all ionizable organic substances of not too low molecular weight, such as bile salts, biocolloids, gums, thymonucleates, more concentrated solutions of sodium penicillin, and a few inorganic examples such as the soluble silicates. Bolam [56] has carried out studies with derivatives of phenanthrene sulfonic acids, first described by Sandquist.

The innumerable water-soluble dyes [57] would deserve more space but for the fact that too few have been submitted to quantitative physical chemical study, such as determination of osmotic coefficient. It has already been emphasized that study of electrolytic conductivity
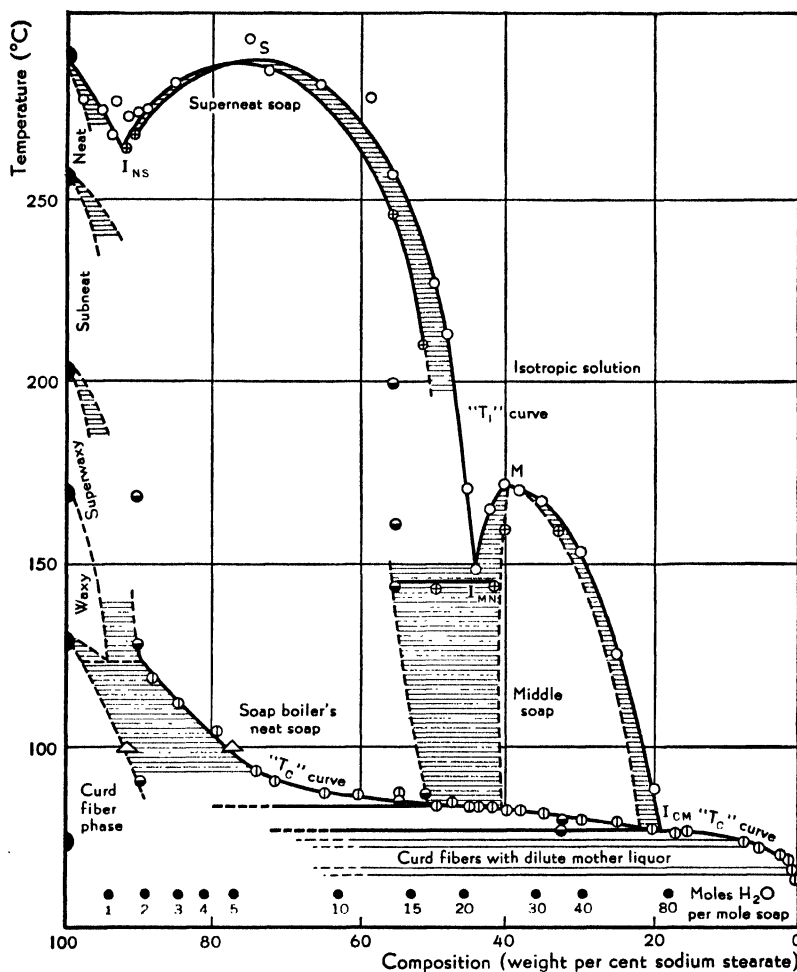
FIGURE 17·23 Typical phase diagram for sodium stearate–water. O, T_l; $\oplus$, T_c; $\bullet$, dilatometric results; $\oplus$, T_o; $\ominus$, other microscopic observations; Δ, vapor pressure results. Taken from McBain, Vold, and Frick, Text Ref. No. 54.

alone may be misleading. The dyes permit a further method of measurement because of their absorption spectra, which change with concentration. For example, Vickerstaff and Lemin [58] conclude that a solution of methylene blue is not a simple mixture of monomer and dimer, but contains higher polymers, especially at high concentration. Some dyes even form jellies. Pauli and Lang [59] have given suggestions as to the form of micelles of Congo red, benzopurpurin, as well as of a few other dyes.

Solubilization and Colloidal Micelles

The colloidal micelles of soaps and detergents have the remarkable property of incorporating in or upon themselves large quantities of molecules which are otherwise insoluble in water. This process leads to a reversible equilibrium.[60]

The contrast between solubilization and ordinary true solution in mixed solvents is illustrated in Figure 17·24. It is seen that a good true solvent such as acetone is rapidly spoiled by small additions of a diluent or nonsolvent such as water, and conversely a large amount of organic solvent has to be added to the water before the solubility of the dye becomes appreciable. In contrast, the solubilization by colloidal electrolytes is already prominent in very dilute solution so that in this respect a dilute solution of colloidal electrolyte can often substitute for a much more concentrated solution of an organic solvent.

Similar phenomena are observed in nonaqueous solutions of colloidal electrolytes. This process is quite unlike emulsification in that the chemical potential of the solubilized molecules is greatly lowered by the solubilization, so that the soap solution will take up unsaturated vapor or unsaturated solution from another phase.[61]

X-ray examination shows that solubilization is usually accompanied by expansion of the micelles. The long spacings increase as is illustrated in the original German work, Figure 17·25, of Hess and collaborators. Here the hydrocarbon lies in thick layers between the hydrocarbon ends of the soap. Quantitative examination shows that not all the solubilization occurs in layers in Hess micelles, but that other types of solubilization occur and that other types of micelles are responsible for some of the solubilization. It is possible also that some molecules interpenetrate between the soap molecules. Those who assume a liquid
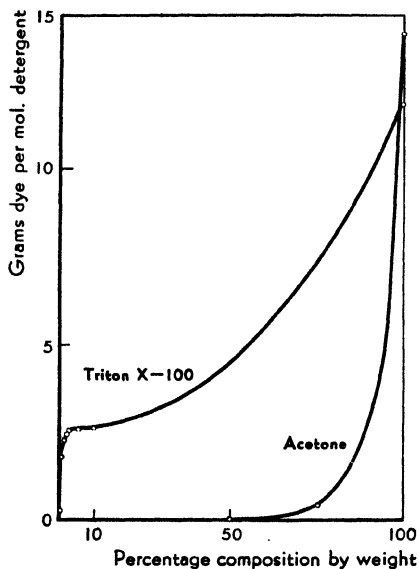
FIGURE 17·24 Solubilization versus solvent action in solutions of Triton X–100 and acetone. Taken from Green and McBain, J. Phys. Chem., 51, 286 (1947).

interior of spherical or cylindrical micelles sometimes assume solution in the hydrocarbon tails; that means irregular juxtaposition of such molecules and tails, since only one molecule can occupy the same space at any one time, and separate molecules always keep about 4–5 Å apart.

Lester Smith [62] suggested that solubilization could be effected by adsorption on the exterior of the colloidal particles. A clear case where solubilization cannot be due to layering within a micelle or solution with the micelle is provided by the solubilization of dimethylphthalate,[63] which is insoluble in water and in hydrocarbons and yet is freely solubilized by potassium laurate without expansion of the x-ray spacings. Added salts, which normally promote formation of lamellar micelles and greatly increase the solubilization of hydrocarbons, depress that of dimethylphthalate, and the amount of dimethylphthalate solubilized per molecule of soap is much less in concentrated solution of soap. Some detergents

FIGURE 17·25 Diagram showing benzene solubilized by a solution of sodium oleate with the benzene layer between the hydrocarbon ends of the soap. Taken from Hess, *Fette und Seifen*, **46**, 572 (1939).

do not form lamellar micelles in any concentration; others do in more concentrated solutions, whereas in more dilute solutions, small micelles preponderate, such as the McBain ionic micelles or fragments or nuclei of lamellar micelles, and it is these that solubilize dimethylphthalate, as is shown in Figure 17·26.

An alternative method of solubilization of polar compounds, such as alcohols or amines of medium length, is by *interpenetration*, as in Figure 17·26a, giving a mixed lamellar micelle.

Another example where solubilization per molecule of detergent is greater in moderate dilution than in more concentrated solution is given in Figure 17·27. Another striking example is that found by Lissant.[64] The solubilization of benzene per molecule of dodecylamine hydrochloride is less in N/4 solution than in N/10 solution. All these examples are further confirmation of the existence of different kinds and sizes of micelles in different ranges of concentration of the same colloidal electrolyte.

FIGURE 17-26 Scale diagram of solubilization of dimethylphthal-ate molecules by polar groups of small micelles. Taken from McBain and McHan, *J. Am. Chem. Soc.*, **70**, 3838 (1948).
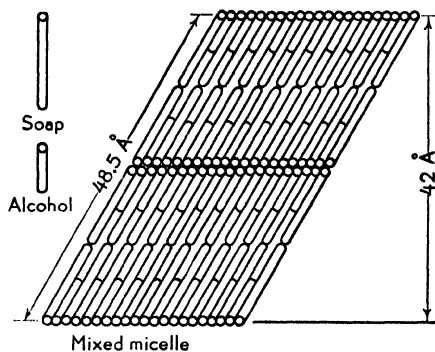
Mixed micelle

FIGURE 17-26a A mixed, or interpenetrated, lamellar micelle of potassium oleate solubilizing an equimolecular amount of *p*-methyl cyclohexanol. After Schulman and Riley, Text Ref. No. 70.

FIGURE 17-27 Solubilization of Orange OT by zephiran at 25° C. Taken from McHan, H., Master's Thesis, Stanford University (1947).
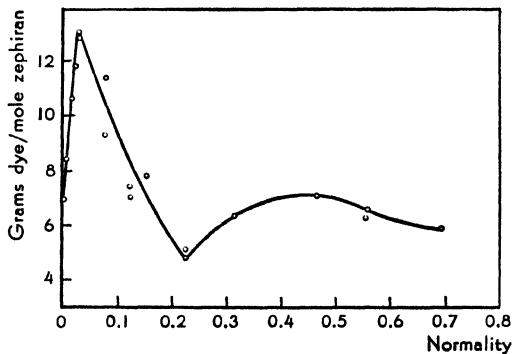
That solubilization goes parallel with colloid formation has been shown by a variety of methods. It is illustrated in the comparison of the osmotic coefficients g and the solubilization of dye per mole of catol [65] in Figure 17-28. It has been repeatedly shown that solubilization begins distinctly below the ordinary critical concentration, the latter being itself lowered, since the complex between detergent and solubilized molecules is formed with positive affinity. Addition of salts so promotes the formation of colloid that solubilization becomes prominent even in extreme dilution. The effect of salts in promoting solubilization of dye or of chlorophyll or of hydrocarbon is illustrated in Figure 17-29.

Solubilization of water-insoluble hydrocarbons and dyes increases extremely rapidly with increase in the molecular weight of the detergent. For instance, with the potassium soaps where the amount of hydrocarbon chain present increases from the octoate to the myristate only in the proportion 1 : 1.25 : 1.50 : 1.75, the amount of dye solubilized by each molecule of hydrocarbon chain increases quite disproportionately as 1 : 2.14 : 6.48 : 11.61,[66] which is incompatible with the solubilization by solution in those hydrocarbon chains. This conclusion deals with amounts and therefore is quite independent of geometrical considerations.

Simultaneous solubilization of two substances sometimes enhances and sometimes decreases the solubilization of the first, even when one is a solvent for the other.[67]

Solubilization expressed in moles solubilized per mole of detergent is greatest by far for small molecules and decreases rapidly with molecular weight of the substance to be solubilized. Frequently polar substances are more strongly solubilized than nonpolar. Other regularities of structure and relation of solubilized material to detergent have been found.[68]
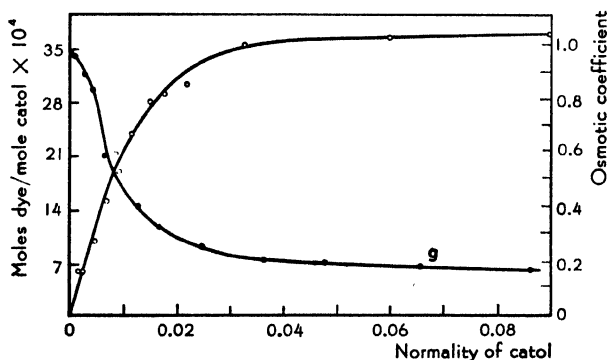
FIGURE 17·28 Solubilization of Orange OT by the cation-active detergent catol as compared with the amount of colloid, indicated by the departure of the osmotic coefficient from the value of unity. Taken from McBain, Wilder, and Merrill, Text Ref. No. 65.

FIGURE 17·29 Effect of salts on solubilization. Solubilization of Orange OT by solutions of a cation-active detergent with and without added potassium chloride. Taken from McBain, Wilder, and Merrill, Text Ref. No. 65.
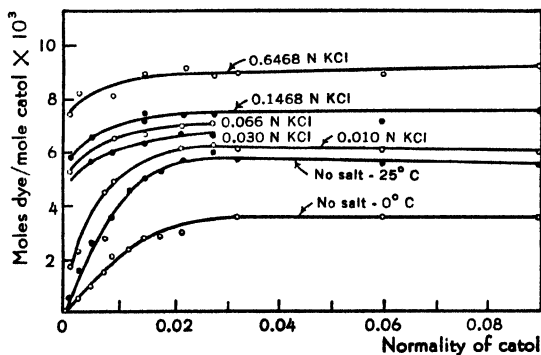
A long series of pure colloidal electrolytes and of commercial detergents have been compared for their power of solubilizing dyes. It proved to be of great interest that the loci for the initiation of many polymerization reactions, such as the formation of synthetic rubber, occur in the hydrocarbon solubilized within the micelles of the soap or detergent used as emulsifying agent.[69]

Solubilization, Hydrotropy, and Emulsions

Hydrotropy occurs in concentrated solutions of salts or of colloidal electrolytes. Not all instances involve colloids, but some do. Some cases involve a particular kind of solubilization that makes a continuous bridge to emulsions. These cases have been examined by Schulman and his collaborators.[70] They began with transparent oil systems containing soap with an at least equimolecular amount of an amphiphilic molecule such as an aliphatic alcohol or amine or a cresol. To these nonconducting solutions water may be progressively added in amounts equal to the oil without producing conductivity, but only developing some Tyndall scattering of light. Large quantities of water result in emulsions. When they began with fairly strong aqueous soap solutions containing at least an equimolecular amount of the amphiphilic substance, a hydrocarbon could be added in large quantities, much larger than in ordinary solubilization.

The concept explaining both these cases is a mixed monolayer of soap and amphiphilic alcohol, interpenetrating, giving a fluid or flexible protective film. The ordinary lamellar micelle is not flexible because of the parallel close packing of the soap molecules. However, the interpenetration by alternate amphiphile molecules plasticizes it so that it can bend round and enclose a sphere of oil or water. Thus arise the hydrophilic oleomicelle and the oleophilic hydromicelle, Figure 17·30, the only completely satisfactory model of a spherical micelle and one whose existence is supported by x-ray evidence. Such micelles must have a total surface just sufficient to accommodate the mixed monolayer of soap and amphiphile, so that as more and more enclosed liquid is added, the droplets must increase in size and decrease in number until a visible, stable emulsion is formed. It is seen that such a model avoids the unlikely assumption that solubilized hydrocarbons can be dissolved in oriented hydrocarbon tails of soap molecules. A similar flexibility of a pure soap monolayer can be attained by heating it above the melting point of the soap — above 50° for ethanolamine oleate.[71]

It is significant that the solubilized "oil" in the lamellar micelles or in the hydrotropic spherical micelles has a lower vapor pressure than the pure oil. Hence the oil is not like bulk liquid but has modified properties such as are discussed in Chapter 4. Emulsion droplets, on the other hand, have a vapor pressure at least as great as ordinary liquid.

| Hydrophilic oleomicelle | Oleophilic hydromicelle |
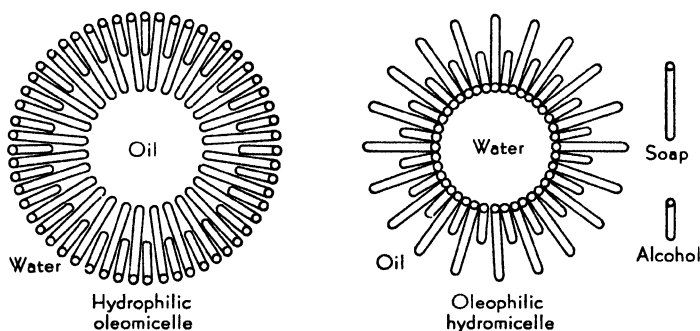
FIGURE 17·30 Spherical micelles in hydrotropic systems where the interfacial monolayer consists of all the soap and an equimolecular amount of an interpenetrating amphiphile such as an alcohol. These can be enlarged by adding further enclosed liquid to produce coarser and coarser droplets up to emulsion size.

WIDESPREAD OCCURRENCE AND APPLICATIONS

Solubilization is a phenomenon of widespread importance. It supplies a means of bringing into solution any substance in any desired solvent provided there is a suitable detergent or colloidal electrolyte available. It finds numerous applications in industry. Emulsion polymerization has already been mentioned. Many dyeing processes would be impossible without it. It is a factor in detergency. Even in non-aqueous media, such as oils, a detergent can solubilize sludge and waste products, as for example in an engine. By its means essential oils and perfumes are held. Dyes and other materials can be carried through membranes even though the pores are too small to permit passage of anything larger than molecules, providing there is a solvent or carrier on the other side. Water-insoluble substances of great physiological importance, for instance vitamins A and K, are carried in the body fluids in solubilized form. The various proteins are colloidal electrolytes and in addition to their many other important characteristics have this solubilizing power. Dyes are eliminated in the bile, for most bile salts are typical colloidal electrolytes. These may serve to indicate the many diversified instances of solubilization in industry and in vital processes.

Detergency

A discussion of soaps and synthetic detergents must include at least a reference to detergency. The primary requisite for detergent action is that the detergent solution should wet the soiled surface and should then loosen the soil. This process has actually been observed under the microscope. The soil is usually a mixture of oily material and solid particles, such as silicious and carbonaceous matter. Oiled iron silicate has sometimes been used.[72]

Wetting usually goes parallel with low surface tension and, in particular, with low interfacial tension. The characteristic surface tension curves for detergents, whether anionic or nonionic, exhibit a very great lowering in dilute solutions, say 0.1%, the surface tension of water being lowered from about 72 dynes to between 25 and 36 dynes, and thereafter remaining practically constant with increase of concentration, first passing through a shallow minimum which in some cases has been proved to be due to impurities. Interfacial tension sometimes goes down progressively to a negligible value. It is interesting to note that an appearance of complete wetting does not always mean that the dirt has been removed. For example, glass soiled with oleic acid and rinsed with dilute sodium oleate solution shows a zero receding angle of contact; but if the surface is tested with water, it is easy to demonstrate by means of the advancing contact angle that the surface is still contaminated with oleic acid.[73]

The usual practical test for penetration is either that of Draves-Clarkson,[74] which depends upon the time required for a skein of cotton to be completely wetted and sink in a solution of detergent, or the thick canvas disc method of Seyforth and Morgan.[75] Cupples [76] has suggested a visual test as to whether a drop of a detergent solution spreads on an oily piece of glass or on a paraffined filter paper.

Spring [77] studied the removal of dirt that was free from all fatty or oily matter and emphasized the importance of protective action of the soap upon the displaced particles of soil.

Mechanical working is a prime factor in assisting detergent action and loosening the dirt.

Basing it upon additional investigations, the writer set up the typical equation for detergency:

$$\text{fabric} \cdot \text{dirt} + \text{soap} \rightarrow \text{fabric} \cdot \text{soap} + \text{dirt} \cdot \text{soap}$$

This emphasizes that the dirt, which is sorbed on or attached to the fabric, is replaced by some form of soap, acid or alkaline, and then the

dirt itself has to be held so effectively in the detergent solution that it is not redeposited upon a cleaner portion of the fabric. Hence the emulsifying or protective power of the detergent is an essential factor in a good detergent. This is tested by measurements of emulsification, and the stability of the emulsion is determined in a centrifuge.[78] Absence of redeposition of dirt has been studied under practical laundry conditions by Carter.[79] Most writers regard studies using the standardized Launderometer as the most convincing.

The most important improvement in ordinary alkali soap would be to insure that any lime soap formed in hard water would be kept in dispersion and therefore carried away in rinsing. The germicidal action of ordinary soaps as distinguished from some synthetic detergents lies almost entirely in the bacteria being carried away in the rinsing.

Water-soluble soil does not need any detergent, and mere rinsing with water suffices unless sorption of the solute has been accompanied by hydrolysis or double decomposition. In that case an ion exchange is necessary and the detergent provides the requisite ions; these can, however, be substituted by other electrolytes. It is common practice to add electrolytes to detergents and soap powders as "builders." These sometimes act by promoting the formation of colloid and sometimes by regulating pH. Nonionic detergents have the great advantage of being independent of pH, but they do not supply material for base exchange. Ionic detergents and electrolytes such as free alkali provide ions which help to suspend protected or unprotected colloidal particles since the gegenions cannot be separated from the colloidal particles which they surround.

An earlier historical discussion by the writer [80] was summed up in the following words:

> ... detergent action of soap involved at least five factors: first, the necessity of having the soap in solution; second, power of emulsification, which goes parallel with low surface tension and the formation of surface films; third, wetting power which, like the last, is ascribable to the undecomposed soap itself and brings about penetration of the fabric and also detachment of grease therefrom; fourth, the formation of non-adhesive colloidal sorption compounds with tissue and impurities due sometimes to acid soap, but more often to soap itself, and capable of remaining in stable suspension; and, fifth, the colloidal state of soap in solution.

To this summary the following comments might now be added.

Foaming and sudsing have no direct relation to detergency although certain advertising has conditioned the housewife in some countries to depend upon this as a criterion. For example, dodecyldiethylene triamine is a good detergent that does not lather, and during World

War II a number of nonfoaming detergents were developed in Germany.

It has been shown by Britz in the author's laboratory that with soaps and other detergents maximum sudsing occurs at or near the concentration showing a minimum of surface tension, near the critical concentration, near the point where the emulsion is not yet very stable, and where the concentration is still insufficient for suspension of carbon black.

REFERENCES

Colloidal Electrolytes

"A General Discussion of Colloidal Electrolytes," *Trans. Farad. Soc.*, **31**, 1 421 (1935)

Text References

1. MCBAIN, J. W., CORNISH, E. C. V., and BOWDEN, R. C., *Trans. Chem. Soc.* (London), **101**, 2042 (1912)
2. LAMBERT, J. M., *J. Coll. Sci.*, **2**, 479 (1947); also EPTON, S. R., *Nature*, **160**, 795 (1947)
3. NORRIS, M. H., *J. Chem. Soc.*, **121**, 2161 (1922)
4. MCBAIN, J. W., VOLD, R. D. and VOLD, M. J., *J. Am. Chem. Soc.*, **60**, 1866 (1938)
5. MCBAIN, J. W., LAURENT, P. and JOHN, L. M., *J. Am. Oil Chem. Soc.*, **25**, 77 (1948)
6. MCBAIN, LAURENT and JOHN, *loc. cit.*
7. MCBAIN, J. W. and TAYLOR, M., *Z. physikal. Chem.*, **6**, 179 (1911)
   MCBAIN, J. W., CORNISH and BOWDEN, *loc. cit.*
   MCBAIN, J. W., *Trans. Faraday Soc.*, **9**, 99 (1913); *Kolloid Z.*, **12**, 256 (1913)
   BUNBURY, H. M. and MARTIN, H. E., *Trans. Chem. Soc.* (London), **105**, 417 (1914)
   MCBAIN, J. W., LAING, M. E. and TITLEY, A. F., *ibid.*, **115**, 1279 (1919)
8. MCBAIN, J. W. and SALMON, C. S., *J. Am. Chem. Soc.*, **42**, 426 (1920)
9. MCBAIN, J. W. and SALMON, C. S., *Proc. Roy. Soc.*, **97A**, 44 (1920)
   MCBAIN, LAING and TITLEY, *loc. cit.*
10. MCBAIN and TAYLOR, *loc. cit.*
    BUNBURY and MARTIN, *loc. cit.*
    MCBAIN, LAING and TITLEY, *loc. cit.*
    MCBAIN and SALMON, *loc. cit.*
11. MCBAIN, M. E. L., DYE, W. B. and JOHNSTON, S. M., *J. Am. Chem. Soc.*, **61**, 3210 (1939)
12. WRIGHT, K. A., ABBOTT, A. D., SIVERTZ, V. and TARTAR, H. V., *J. Am. Chem. Soc.*, **61**, 549, 552 (1939)
13. LOTTERMOSER, A. and PÜSCHEL, F., *Kolloid-Z.*, **63**, 175 (1933)
14. REED, R. M. and TARTAR, H. V., *J. Am. Chem. Soc.*, **58**, 322 (1936); PAQUETTE, R. G., LINGAFELTER, E. C. and TARTAR, H. V., *ibid.*, **65**, 686 (1943); SCOTT, A. B. and TARTAR, H. V., *ibid.*, **65**, 692 (1943); SCOTT, A. B., TARTAR, H. V. and LINGAFELTER, E. C., *ibid.*, **65**, 698 (1943)
15. RALSTON, A. W. and HOERR, C. W., *J. Am. Chem. Soc.*, **64**, 97, 772 (1942)
16. WALTON, H. F., HIEBERT, E. N. and SHOLTES, E. H., *J. Coll. Sci.*, **1**, 385 (1946)
17. BUNBURY and MARTIN, *loc. cit.*
    MCBAIN, LAING, and TITLEY, *loc. cit.*
18. GONICK, E., *J. Am. Chem. Soc.*, **67**, 1191 (1945)
19. MCBAIN and SALMON, *loc. cit.*
20. MCBAIN, J. W. and BRADY, A. P., *J. Am. Chem. Soc.*, **65**, 2072 (1943)

21. HARTLEY, G. S., COLLIE, B. and SAMIS, C. S., *Trans. Farad. Soc.*, 32, 796 (1936)
22. HARTLEY, G. S., *J. Chem. Soc.*, 1968 (1938)
    ✸ MERRILL, R. C. and MCBAIN, J. W., *J. Phys. Chem.*, 46, 10 (1942)
23. KLEVENS, H. B., *22nd Annual Colloid Symp.*, M.I.T. (June, 1948)
24. LAING, M. E., *Trans. Chem. Soc.* (London), 113, 435 (1918)
25. HOWELL, O. R. and ROBINSON, H. G. B., *Proc. Roy. Soc.*, 155A, 386 (1936)
    WARD, A. F. H., *Proc. Roy. Soc.*, 176A, 412 (1940)
    EVERS, E. C. and KRAUS, C. A., *J. Am. Chem. Soc.*, 70, 3049 (1948)
    GRIEGER, P. F. and KRAUS, C. A., *J. Am. Chem. Soc.*, 70, 97 (1949)
26. GONICK, E., *J. Phys. Chem.* 50, 291 (1946)
27. MCBAIN, M. E. LAING, *J. Phys. Chem.*, 47, 196 (1943); see also NEFF, L. L., WHEELER,
    O. L., TARTAR, H. V. and LINGAFELTER, E. C., *J. Am. Chem. Soc.*, 70, 1989 (1948)
28. LAING, M. E., *J. Phys. Chem.*, 28, 673 (1924)
29. MCBAIN, J. W. and JENKINS, W. J., *J. Chem. Soc.*, 121, 2325 (1922)
30. MCBAIN, *J. Am. Chem. Soc.*, 50, 1636 (1928)
    MCBAIN, J. W. and BETZ, M. D., *J. Am. Chem. Soc.*, 57, 1905 (1935)
    MCBAIN, J. W. and SEARLES, J., *J. Phys. Chem.*, 40, 493 (1936)
    CORRIN, M. L. and HARKINS, W. D., *J. Am. Chem. Soc.*, 69, 683 (1947)
31. MCBAIN and SEARLES, *loc. cit.*
32. HARKINS, W. D., *Science*, 105, 635 (1947); MATTOON, R. W., STEARNS, R. S. and
    HARKINS, W. D., *J. Chem. Phys.*, 15, 209 (1947)
33. MCBAIN, J. W., *Kolloid-Z.*, 12, 256 (1913); *Trans. Farad. Soc.*, 9, 99 (1913)
34. HOWELL and ROBINSON, *loc. cit.*
35. RALSTON, A. W., EGGENBERGER, D. N. and DUBROW, P. L., *J. Am. Chem. Soc.*, 70,
    977 (1948)
36. HARTLEY, COLLIE, and SAMIS, *loc. cit.*
37. BROWN, G. L., GRIEGER, P. F. and KRAUS, *J. Amer. Chem. Soc.*, 71, 97 (1949)
38. ADAM, N. K., *J. Phys. Chem.*, 29, 98 (1925)
39. STAUFF, J., *Kolloid-Z.*, 96, 244 (1941)
40. MEYER, K. H. and VAN DER WYK, A., *Helv. chim. Acta*, 20, 1321 (1937)
41. HESS, K. and GUNDERMANN, J., *Ber.*, 70, 1800 (1937)
42. HESS, K., KIESSIG, H. and PHILIPPOFF, W., *Naturwissen.*, 26, 184 (1938)
    KIESSIG, H. and PHILIPPOFF, W., *ibid.*, 27, 593 (1939)
    HESS, K., PHILIPPOFF, W. and KIESSIG, H., *Kolloid-Z.*, 88, 40 (1939)
    HESS, K., *Fette und Seifen*, 46, 572 (1939)
    KIESSIG, H., *Kolloid-Z.*, 96, 252 (1941)
    PHILIPPOFF, W., *Kolloid-Z.*, 96, 255 (1941)
    HESS, K., KIESSIG, W., and PHILIPPOFF, W., *Fette und Seifen*, 48, 377 (1941)
    KIESSIG, H., *Kolloid-chem. Taschenbuch*, A. Kuhn, Leipzig (1944 and 1948), pp.
    154–161
    HESS, K., *Fette und Seifen*, 49, 81 (1942)
    KIESSIG, H., *Kolloid-Z.*, 98, 213 (1942)
43. MCBAIN, J. W., KAWAKAMI, Y. and LUCAS, H. P., *J. Am. Chem. Soc.*, 55, 2762 (1933)
44. STAUFF, J., *Kolloid-Z.*, 89, 224 (1939); 96, 244 (1941); *Naturwissen.*, 27, 213 (1939)
45. VETTER, R. J., *J. Phys. Coll. Chem.*, 51, 271 (1947)
46. LAWRENCE, A. S. C., *Trans. Farad. Soc.*, 34, 660 (1938)
47. GONICK, E., *J. Coll. Sci.*, 1, 393 (1946)
    MCBAIN, J. W. and WORKING, E. B., *J. Phys. Coll. Chem.*, 51, 974 (1947)
48. HARTLEY, G. S., *Aqueous Solutions of Paraffin Chain Salts*, Actualités Scientifiques
    et Industrielles #387, Hermann et Cie, Paris (1936); *Kolloid-Z.*, 88, 22 (1939)
49. WARREN, B. E., *Phys. Rev.*, 44, 969 (1933)
50. WALTON, HIEBERT, and SHOLTES, *loc. cit.*
51. KLEVENS, H. B., *J. Am. Chem. Soc.*, 70, 3526 (1948)

52. HARKINS, W. D., *J. Chem. Phys.*, **16**, 156 (1948)
53. MCBAIN, J. W. and MARSDEN, S. S., *Acta Crystallograph.*, **1**, 270 (1948); *J. Am. Chem. Soc.*, **70**, 1973 (1948); *J. Chem. Phys.*, **16**, 633 (1948)
54. MCBAIN, J. W., VOLD, R. D. and FRICK, MARY, *J. Phys. Chem.*, **44**, 1013 (1940)
55. MARSDEN, S. S. and MCBAIN, J. W., *loc. cit.*
56. BOLAM, T. R. and HOPE, J., *J. Chem. Soc.*, 843, 850 (1941)
57. VALKÓ, E., *Textilchemie; Kolloidchemische Grundlagen der Textilveredlung*, Julius Springer, Berlin (1937) pp. 319–488
58. VICKERSTAFF, T. and LEMIN, D. R., *Nature*, **157**, 373 (1946)
59. PAULI, W. and LANG, F., *Monatsh.*, **67**, 159 (1936)
60. MCBAIN, J. W. and MCBAIN, M. E. L., *J. Am. Chem. Soc.*, **58**, 2610 (1936)
61. MCBAIN, J. W. and O'CONNOR, J. J., *J. Am. Chem. Soc.*, **62**, 2855 (1940); **63**, 875 (1941)
62. SMITH, E. LESTER, *J. Phys. Chem.*, **36**, 1401, 1672, 2455 (1932)
63. MCBAIN, J. W. and MCHAN, H., *J. Am. Chem. Soc.*, **70**, 3838 (1948)
64. LISSANT, K. J., Ph.D. Thesis, Stanford University, 1947
65. MCBAIN, J. W., WILDER, A. G. and MERRILL, R. C., *J. Phys. Coll. Chem.*, **52**, 12 (1948)
66. MCBAIN, J. W. and GREEN, A. A., *J. Am. Chem. Soc.*, **68**, 1731 (1946)
67. GREEN, A. A. and MCBAIN, J. W., *J. Phys. Coll. Chem.*, **51**, 286 (1947)
68. MCBAIN, J. W. and RICHARDS, P. H., *Ind. Eng. Chem.*, **38**, 642 (1946)
69. HARKINS, W. D., *J. Chem. Phys.*, **13**, 381 (1945); STANDINGER, J. J. P., *Chem. and Ind.*, Sept. 4, 1948, 563
70. SCHULMAN, J. H.; with HOAR, T. P., *Nature*, **152**, 102 (1943); with MCROBERTS, T. S., *Trans. Faraday Soc.*, **42B**, 165 (1946); with RILEY, D. P., *J. Coll. Sci.*, **3**, 383 (1948)
71. PINK, R. C., *Trans. Faraday Soc.*, **42B**, 170 (1946)
72. SNELL, C. T., *J. Chem. Ed.*, **24**, 505 (1947)
73. Cf. LIDDIARD, P. D., *Chem. & Ind.*, **28**, 435 (1948)
74. DRAVES, C. Z., *Am. Dyestuff Reptr.*, **28**, 421, 425 (1939)
75. SEYFORTH, H. and MORGAN, O. M., *Am. Dyestuff Reptr.*, **27**, 525 (1938)
76. CUPPLES, H. L., *Ind. Eng. Chem.*, **27**, 1219 (1935)
77. SPRING, W., *Kolloid-Z.*, **4**, 161 (1909); **6**, 11, 109, 164; *Arch. sci. phys. et nat.*, **28**, 569; **29**, 41; **36**, 80; *Rec. trav. chim.*, **29**, 1 (1910); *Bull. acad. roy. med. Belg.*, 1909, 187, 949; 1911, 24, 17
78. MERRILL, R. C., *Ind. Eng. Chem.* (Anal. Ed.), **15**, 743 (1943)
79. CARTER, J. D., *Ind. Eng. Chem.*, **23**, 1389 (1931)
80. MCBAIN, J. W., in *Advances in Colloid Science*, Vol. I, Interscience Publishers, Inc., New York (1942), p. 103

SOLUTIONS OF PROTEINS

Proteins occupy a key position in life processes, being the most important constituent of protoplasm, and it is not surprising that they should be more intensively studied than any other class of colloids. For detailed accounts references should be made to one of the monographs on the subject [1] or to the symposia that are held at frequent intervals.[2]

Proteins are amphoteric substances derived essentially from the combination of amino acids and characterized by the presence of the polypeptide chain made up of repetitions of the peptide unit:

$$
\begin{array}{c}
R \\
| \quad \diagdown H \\
C \\
\diagup \quad \diagdown \quad \diagup \\
C \qquad N \\
\| \qquad | \\
O \qquad H
\end{array}
$$
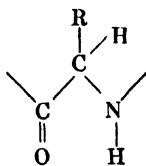
where each different R is a side chain characteristic of particular amino acids to be found in various parts of the chain. Some of these side chains carry carboxyl and amino groups. Others are paraffinic, alcoholic or phenolic, or SH, S—S, or S—CH_3 groups. Furthermore, carbohydrates, nucleic acids, and other simple substances are found widely associated with the proteins, held weakly or strongly. Such entities as lipoproteins, glucoproteins, and nucleoproteins often form the functional unit in biological systems, and their occurrence is largely responsible for the extreme variations in the physical characteristics of different proteins and for the conditions necessary for their isolation.

Since the α-amino acids can unite either with strong acids or with strong bases, there is an intermediate value of acidity or pH where the number of positive charges and of negative charges on the protein are equal. This is the isoelectric point. Since the positive and negative charges are on different parts of the molecule, the protein particle forms a zwitterion. However, many of the amino and carboxyl groups are not free to react, and even for a protein such as egg albumin with molecular weight 35,000–45,000 (1944) at the point of maximum dissociation, there will not be more than about 28 charges. Eilers [3] has recently shown that many of the properties of proteins in solution, such as their titration curves, are merely the superimposed values for the different constituent amino acids.

It is interesting to inquire how far apart these charges would be if they are equally spaced on the exterior of a spherical surface of the radius ascribed to egg albumin, 22 Å. The result is to show that they would be 15 Å apart; that is, they would be as far apart as the ions in any $N/2$ solution of a uni-univalent electrolyte such as potassium chloride. Hence, the ionic atmospheres corresponding to the separate charges are largely independent and resemble those of a univalent electrolyte in contrast to an ordinary polyvalent ion which has all its like charges concentrated in a point with correspondingly enhanced ionic strength.

At the isoelectric point many properties of the protein solution are at either a maximum or a minimum. Among the latter are the viscosity, osmotic effects such as osmotic pressure, electrical conductivity, diffusion, electrokinetic movement, solubility, and swelling; while the coagulability by heat or by alcohol and the ability to spread on water, and the foaming power, are at a maximum. It may be noted that when a protein spreads to a close-packed film on water of suitable pH, the layer is 10–20 Å deep, which represents the average length of the side chains in the amino acids. Clearly it is not the tendency for these like groups to get together that brings them out upon the surface, but rather the expulsion of these insoluble parts of the protein molecule from the water in order that the water molecules may come into contact with each other. The sheet of protein is held to the water surface by its soluble or ionizable points.

Figure 18·1 illustrates [4] experiments of Gorter and collaborators which show the dependence of the area of spreading of egg albumin upon pH of the underlying solution, which clearly shows the maximum spreading at the isoelectric point 4.8. Incidentally, this area for most proteins is roughly 1 mg/sq meter, just as it is for many dyes and other

organic substances. The surface activity of proteins has been extensively reviewed by Neurath and Bull [5] and that of egg albumin in particular more recently by Trurnit.[6]

The surface areas of proteins in various states of preparation as determined by the nitrogen method, Chapter 3, range from 0.2 to 14 sq m/gm. The areas obtained from water sorption are from 40- to 200-fold larger than the nitrogen areas, leading to the conclusion that water is more closely related to hydration of specific groups.[7]
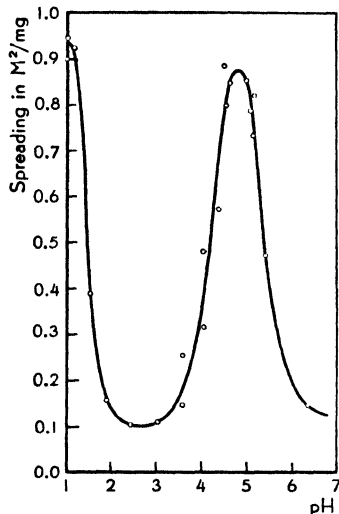
FIGURE 18·1 Diagram showing how the area of spreading of egg albumin depends upon pH of the underlying solution. Taken from Gorter et al., Text Ref. No. 4.

FIGURE 18·2 Diagram showing the effects of progressive amounts of hydrochloric acid added to a protein solution.
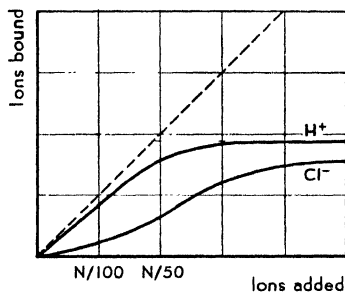
If now we add progressive amounts of hydrochloric acid to the protein solution as Pauli did with serum albumin,[8] we obtain the result represented by the diagram in Figure 18·2. The dashed line running at an angle of 45° represents what would have been observed if every ion added had become attached to the protein whether by chemical action or by sorption. Any hydrogen ions added which did not so combine were detected by a hydrogen or glass electrode, and corresponding electromotive force measurements determined the free or unattached chloride ions. It is seen that the maximum difference between the two, that is, the maximum ionization, is attained in N/50 hydrochloric acid. Here the behavior is that of a stable colloidal electrolyte. Those properties which we have enumerated as being at a minimum at the isoelectric point are here at a maximum, and vice versa. At a higher pH than the isoelectric point proteins combine with alkalies and bind cations, whereas below the isoelectric point they combine with acids and take up only anions.

When anionic detergents such as long chain sodium sulphates are added to solutions of proteins, they form adsorption complexes. The solubility of these complexes decreases to a minimum as the detergent protein ratio is increased and then rises again as still more detergent is added. Between certain detergent protein ratios, the complexes separate as coacervates.[9]

It is possible to obtain minimum values for the particle or molecular weight of proteins from their combining weight, as for example with carbon monoxide. Minimum values again may be obtained from the irreducible value of some element or component after exhaustive purification. Hemoglobin from the cow combines with one mole of carbon monoxide per 16,700 gm and the same weight of hemoglobin contains 1 gm atom of iron. A similar iron content is observed with hemoglobin from the dog, horse, and hen, with somewhat lower weights for the cat and especially the hog. Similarly the content of sulfur or sulfide or the amino acid arginine has been employed. A satisfactory multiple of all these values is 67,000. This is commonly taken as the molecular or particle weight, because it has been obtained by means of the ultracentrifuge,[10] osmotic pressure,[11] and diffusion.[12,13] Cohn's best value is 68,000. There is still uncertainty as to whether this represents a true molecular weight or that of a reproducible aggregate, especially since proteins may be split into fractions sometimes 2, 4, or 8 times smaller than the original just by changing the pH too much or by adding such compounds as urea and clupein. Stronger acid produces still smaller entities. Frequently this splitting is reversible. Even the fractions are held by some authorities [14] to be built up from fairly small subunits.

Denaturation of Proteins

Denaturation is an important alteration in the properties of proteins which tends to make them insoluble in water, transforming a hydrophile to a hydrophobe sol. The insolubility is especially noticeable at the isoelectric point. Denaturation produces increased opalescence and has been shown to be accompanied by diminished hydration of the protein.

Denaturation involves two separate processes, the first being the lowering of solubility and the second the readiness of coagulation. This is clearly shown when albumin is boiled in dilute acid or in the presence of a thiocyanate. Flocculation does not occur until the acid is neutralized by alkali or until the acid or salt has been removed by dialysis or electrodialysis.

The two chief causes of denaturation are heating and adsorption. The effect of heat is of great importance because it sets a limit to the existence of life. An organism dies when its protein denatures. It has been shown that living organisms if sufficiently desiccated may continue to live at temperatures even exceeding the boiling point of water. Heat denaturation is accompanied by a large evolution of heat. It is strongly accelerated or catalyzed by both hydrogen and hydroxyl ions. Hence we have the interesting result that the minimum rate of denaturation has no reference to the isoelectric point of the protein but comes at the neutral point of water, whereas the maximum rate of coagulation is at the isoelectric point.

Denaturation is for some proteins wholly or partially reversible. This is true of silk fibroin, as discussed in Chapter 24. For others, such as egg albumin, it is irreversible except by some indirect path or chemical treatment. Salts, as will be seen, produce complicated effects. A few like thiocyanate, salicylate, and benzoate appear to prevent denaturation.

The other chief cause of denaturation is spreading on surfaces, or adsorption. This is familiar in the head of foam on beer. Egg albumin is so thoroughly denatured by adsorption that if a solution is allowed to fall a distance of a meter in the form of a fine stream or spray, the albumin collects as a denatured scum, leaving the water almost quantitatively free from protein. Similarly, egg albumin in very dilute solution collects as an insoluble pellicle on the surface.

Often an already denatured protein will not spread, but this property may be restored by mild treatment with an enzyme, such as pepsin.[15] Denaturation is ascribed to the rupture of certain linkages in the native protein and probably involves partial combination between the ionizable groups of the protein. It may possibly include more thoroughgoing combinations, as through such linkages as S—S, CO—S, NH—S, or CO—NH—CO. It is accompanied by the appearance of thiol and other reducing groups.[16] Denaturation results in a great elongation or uncoiling of the original protein molecule.

Salting-out and Coagulation of Proteins

Gelatin may be contrasted with albumin. Gelatin, like agar, or sericin, is not denatured by heating. Instead, heating hastens the loss of the power to gelatinize until it is quite gone. This change is also accelerated by both hydroxyl and hydrogen ions. In contrast, albumin does not gelatinize.

Albumin and gelatin both salt out reversibly in the presence of high concentrations of the salts of the alkali metals and magnesium. This is frequently referred to either as coagulation or as coacervation. However, closer examination [17] shows that it is merely the formation of two liquid layers, closely resembling the phase behavior of soap solutions, and quite analogous to the salting out of aqueous alcohol into two liquid layers by sufficient addition of potassium carbonate. Mere dilution results in the redissolving to one uniform liquid.

In reversible salting out, the lyotropic series is prominent. This series was fully discussed in Chapter 9 and the original observations [18] of Hofmeister there referred to are now given in Table 19.

TABLE 19

CONCENTRATION OF SALTS IN MOLES PER LITER FOUND NECESSARY BY
HOFMEISTER TO SALT OUT EGG ALBUMIN REVERSIBLY IN
NEUTRAL SOLUTION AT 30–40°

	Li	K	Na	NH_4	Mg
Citrate	—	0.56	0.56	—	—
Tartrate	—	0.75	0.78	—	—
Sulfate	0.78	—	0.80	1.00	1.32
Acetate	—	1.67	1.69	—	—
Chloride	—	3.52	3.62	—	—
Nitrate	—	—	5.42	—	—
Chlorate	—	—	5.52	—	—
Iodide	—	—	*	—	—
Sulfocyanide	—	—	*	—	—

* Insufficiently soluble to salt out

In alkaline solution similar results are obtained, but in weak acid (that is, less than 0.03 N) some of the results are irreversible. Now in more strongly acid solutions Pauli has found that the series is *reversed* and the coagulation is irreversible. The irreversible coagulation or denaturation on heating also shows the effect of the lyotropic series. Thus, the coagulation point might be 60° in the neutral albumin but be raised in neutral or alkaline solution by adding salts of the alkali metals, until, as we have seen, with potassium thiocyanate there is no coagulation even on boiling until the thiocyanate is removed. Here again in acid solution, the effect of ions upon the irreversible heat coagulation is reversed, and the thiocyanates cause coagulation at the lowest temperature.

Irreversible coagulation or precipitation is brought about by most other cations or salts of heavy metals. The effects are very specific; sometimes the coagulum may be dissolved by excess of sol or cation.

Reversible and Irreversible Changes

To sum up these changes, a number are reversible, such as drying and redissolving, as with gums, tannin, or albumin, gelatinization generally on cooling, when it occurs, and salting out by salts of the alkali metals. Irreversible coagulation is brought about by heating, by surfaces, by salts, particularly those of the heavy metals, by non-electrolytes, as for example, with gelatin and formaldehyde or with albumin and alcohol. Other irreversible coagula may be produced by freezing, shaking, fermenting, or by the mutual sorption of colloids. Throughout all these, the effect of the lyotropic series, discussed in Chapter 9, is important.

The Purification of Proteins

Since a single protein may contain one or more dozen different kinds of amino acids and since they are so often associated with each other and with other complexes, the purification and characterization of individual proteins has become a fine art. The commonly used method of separating proteins is by precipitation at the isoelectric point or at definite pH's by salt solutions or mixtures of salt with alcohol or other reagents which cover up the soluble side groups on the polypeptide chain. Such precipitation is usually preceded by extraction with lipoid solvents. A highly specific and therefore characteristic precipitation occurs when a toxin is used to precipitate an antitoxin and the very pure antibody is then recovered. The solubilities of different groups of proteins [19] are sufficiently dependent upon concentration of salts to permit their differentiation and separation by progressive increase in the concentration of the salt. However, this fractionation alone seldom results in a pure preparation. It may be supplemented by picking out different crystal forms, using slight changes in pH.[20]

For more rigorous differentiation and separation, albeit of only very small quantities, the ultracentrifuge,[21] as discussed in Chapter 16, diffusion,[22] Chapter 14, and the electrophoretic apparatus of Tiselius are employed. Dielectric studies may be employed.[23]

In general, it is necessary to use several methods because a protein which appears homogeneous to one type of measurement may be shown to consist of more than one individual by another type of measurement, such as electrophoresis. A pure protein is characterized by a definite solubility, just as in the case of ordinary crystalloids, but few preparations achieve this result.

Ion exchange on resins is sometimes employed. The amino acids resulting from protein hydrolysis may be classified as follows: the dicarboxylic acids are separated from the others by adsorption on a basic resin, and the basic diamino acids collected on a weakly acid (carboxylic) resin through which the more neutral acids pass unchanged; the remaining acids are then removed and concentrated by adsorption on a strongly acid resin, through the reaction:

$$R—SO_3' \ldots H^+ + NH_3^+XCOO' = R—SO_3' \ldots NH_3^+XCOOH$$
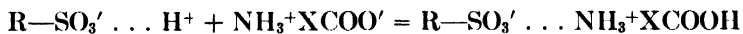
There is some evidence that general adsorption at the resin surface is often superimposed on the exchange reactions at the active groups. Among amino acids tryptophane seems to be preferentially adsorbed, and the effect may be an important one with aromatic compounds in general.

The Electrophoresis of Proteins

The moving boundary method of measuring electrophoresis has long been known, but its refinement, for the purpose of distinguishing between and separating individual proteins, is due to Tiselius.[24] His
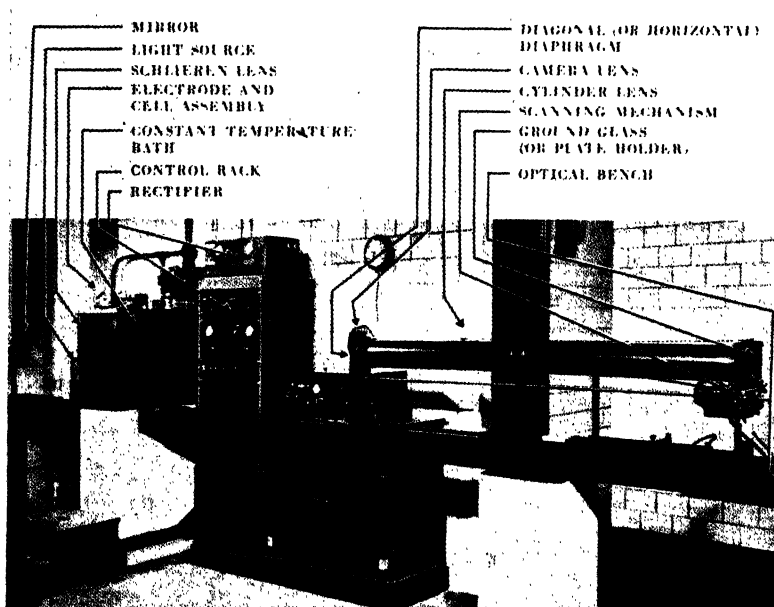
FIGURE 18·3 Modern electrophoresis apparatus, as used in the laboratories of Eli Lilly and Company. From *Research Today*, 1948, through the courtesy of Eli Lilly and Company.

method has been universally adopted and its developments are described by many writers.[25]

The various proteins migrate with different velocities and hence if the electrophoresis is prolonged and the conditions are carefully controlled, a separate boundary develops for each, and the faster ones may even migrate completely away from the slower ones. To minimize volume changes the apparatus is often held near 4° C and stringent requirements have to be met in the solutions surrounding the reversible
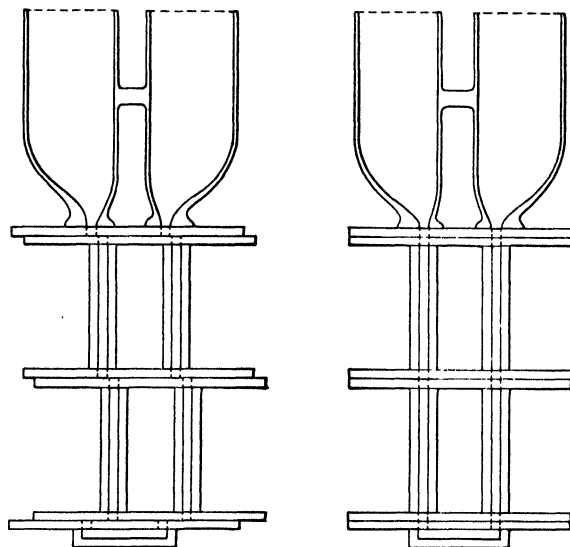
FIGURE 18·4 Cross section of middle portion of U-tube of electrophoresis equipment; *right*, all sections shown in starting position; *left*, shows how portions of the cell may move sidewise, thus dividing the cell into five sections. Taken from Tiselius, A., *The Harvey Lectures*, 1939–40, **XXV**, pp. 37–40.

electrolytes so that the advancing and receding boundaries may be properly developed without modification or distortion. To facilitate observation and measurements the cross section of the electrophoresis cell is made rectangular. If ultraviolet light is to be used, since it only is absorbed by the proteins, these cells are made of fused silica. Alternatively, the refractive indices methods described in the chapter on ultracentrifugation (16) may be employed.

A complete, modern electrophoresis equipment, as used in the Research Laboratories of Eli Lilly and Company, is given in Figure 18·3. A cross section of the middle portion of the "U-tube," leaving out the electrode and much accessory equipment, is given in Figure 18·4. The horizontal portions are ground glass to which the vertical

rectangular portions are sealed. Advantage is taken of the movement past one another to capture the separate proteins by sliding these portions of the cell sidewise, thus dividing a protein into five separate portions, as shown on the left of the diagram.

The beauty of this technique is illustrated in Figure 18·5, again taken from the work of the Lilly Research Laboratory and showing the position of the boundaries for seven constituents of blood plasma
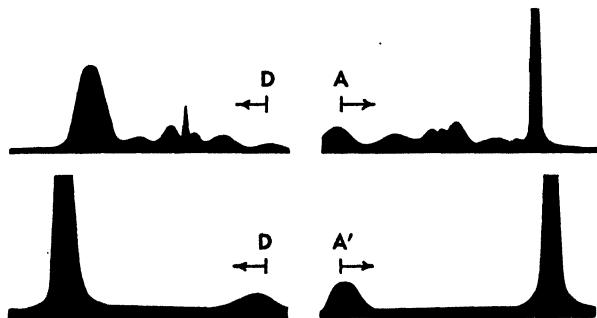
FIGURE 18·5 Patterns showing the resolution of the components of human plasma by the Tiselius electrophoresis apparatus—A, ascending boundaries; D, descending boundaries. The upper two patterns are for plasma; in A the maxima are assigned, from right to left, to albumin, α_1-globulin, α_2-globulin, β_1-globulin, β_2-globulin, fibrinogen, γ-globulin, and the boundary anomaly. . . . The lower two patterns are for albumin. In A′, the peak at the right is albumin; the other corresponds to the boundary anomaly. This shows that the sample is substantially free of all other plasma components. Courtesy of Eli Lilly and Company.

and the resulting pattern for a single pure preparation of albumin. The pattern is a plot of the rate of change of refractive index, that is, the protein concentration against vertical position in the cell. The area under each peak is directly proportional to the concentration of that component. This leads to a quantitative estimate of the proportions of the various proteins present.

The Formation of Antibodies

The most surprising feature with regard to the behavior of proteins and their specificity is the way in which during metabolism and inheritance each protein reproduces further proteins which are either exact replicas or, in the case of antibodies, fit each other like lock and key. Pauling [26] has described a possible process for a particular case of antibodies which are assumed to maintain the same polypeptide chains as the normal globulin, differing only in the way the chain is coiled in the molecule. The mechanism resembles the production of replicas with a template.

REFERENCES

Text References

1. PAULI, WO. and VALKÓ, E., *Kolloidchemie der Eiweisskörper*, Theodor Steinkopff, Dresden and Leipzig (1933); SCHMIDT, C. L. A., *Chemistry of Amino Acids and Proteins*, 2nd ed., Charles C. Thomas, Baltimore (1944); LLOYD, DOROTHY JORDAN and SHORE, AGNES, *Chemistry of the Proteins*, Blakiston's Son and Co., Inc., Philadelphia (1938); COHN, E. J. and EDSALL, J. T., *Proteins, Amino Acids, and Peptides as Ions and Dipolar Ions*, Reinhold Publishing Corp., New York (1943); ANSON, M. L. and EDSALL, JOHN T., editors, *Advances in Protein Chemistry*, Academic Press, Inc., New York, Vol. I (1944), Vol. II (1945), and Vol. III (1947); *Fibrous Proteins*, Society of Dyers and Colourists, Leeds (1946), pp. 1–220

2. For example, *Symposium on Quantitative Biology*, Vol. VI, Biological Laboratory, Cold Springs Harbor (1938); *Proc. Roy. Soc.*, 170A, 40–80 (1939); *Chem. Rev.*, 24, 177–366 (1939)

3. EILERS, H., *Research in Holland*, 1947, 1. See C. A. 42, 8357f (1948)

4. GORTER, E., VAN ORMONDT, J. and DORN, F. P., *Proc. Kon. Akad. Amsterdam*, 35, 838 (1932)

5. NEURATH, H. and BULL, H. B., *Chem. Rev.*, 23, 391 (1938)

6. THURNIT, H. J., *Fortschritte der Chemie organischer Naturstoffe*, Vol. IV, Springer Verlag, Vienna (1945), pp. 415–444

7. BENSON, S. W. and ELLIS, D. A., Chicago meeting of the A.C.S., April, 1948

8. MANABE, K. and MATULA, J., *Biochem. Z.*, 52, 369 (1913)

9. PARKHURST, K. G. A., *Chem. and Ind.*, 1949, 44.

10. SVEDBERG, T., *Ind. Eng. Chem.* (Anal. Ed.), 10, 125 (1938)

11. ADAIR, G. S., *Proc. Roy. Soc.*, 109A, 292 (1925). Similarly PAULI, WO. and FENT, P., *Kolloid-Z.*, 67, 288 (1934), obtained 59,000 for serum albumin carefully freed from electrolyte.

12. NORTHRUP, J. H. and ANSON, M. L., *J. Gen. Physiol.*, 12, 543 (1929). Similarly MCBAIN, J. W., DAWSON, C. R. and BARKER, H. A., *J. Am. Chem. Soc.*, 56, 1021 (1934), obtained 34,000 for egg albumin and KUNITZ, M., ANSON, M. L. and NORTHRUP, J. H., *J. Gen. Physiol.*, 17, 365 (1934), obtained 35,000 for trypsin.

13. BERNAL, J. D., *Nature*, 143, 663 (1939)

14. BERNAL, J. D., *loc. cit.*

15. GORTER, E., *Trans. Farad. Soc.*, 33, 1125 (1937)

16. MIRSKY, A. E. and ANSON, M. L., *J. Gen. Physiol.*, 19, 427 (1936)

17. MCBAIN, J. W. and KELLOGG, F., *J. Gen. Physiol.*, 12, 1 (1928)

18. HOFMEISTER, F., *Arch. expt. Path.*, 25, 13 (1888); 28, 210 (1891)

19. HERRIOTT, R. M., *Chem. Rev.*, 30, 413 (1942)

20. DISTECHE, A., *Nature*, 161, 130 (1948)

21. PICKELS, E. G., *Chem. Rev.*, 30, 341 (1942)

22. LAMM, O. and POLSON, A., *Biochem. J.*, 30, 528 (1936); POLSON, A., *Kolloid-Z.*, 87, 149 (1939); NEURATH, H., *Chem. Rev.*, 30, 357 (1942)

23. ONCLEY, J. L., *Chem. Rev.*, 30, 433 (1942)

24. TISELIUS, A., *Nova Acta Regiae Soc. Sci. Upsaliensis*, VII (4), 1–107 (1930)

25. TISELIUS, A., *The Harvey Lectures*, XXV, 37–40 (1939–40); *Symposium on Electrophoresis*, *Annals of New York Acad. Sci.*, 39, 105–212 (1939); ABRAMSON, H. A., MOYER, L. S. and GORIN, M. H., *Electrophoresis of Proteins and the Chemistry of Cell Surfaces*, Reinhold Publishing Corp., New York (1943); LONGSWORTH, L. G., *Chem. Rev.*, 30, 323 (1942); *Research Today*, 4(2), 78 (1947) Eli Lilly & Company, Indianapolis

26. PAULING, L., *J. Am. Chem. Soc.*, 62, 2643 (1940)

NONAQUEOUS SYSTEMS OF COLLOIDS

Association occurs in nonaqueous media almost as much as in water. Here again there are two types of colloidal particles. The first is the more or less fortuitously formed or unstable colloid, such as the various metal sols that result from the Bredig method of striking an arc between two metal wires immersed in a solvent. Also it may be recalled that many natural and synthetic oriental gems consist of finely divided metallic compounds in a matrix of corundum, Al_2O_3.

The second type is the association colloid, such as the soaps that are soluble in nonaqueous solvents. Direct proof of the association of such soaps is given by the x-ray examination of S. S. Marsden, Jr., in the Stanford laboratory, who obtained with triethanolamine laurate in dry benzene the characteristic spacings of lamellar micelles consisting of soap molecules placed end to end. This is evidently due to the presence of inverted soap micelles with the polar groups outwards, as pictured by Lawrence.[1] Such association colloids as the detergents and colloidal electrolytes that are soluble in the respective solvents can solubilize in them otherwise insoluble substances. For example, various dyes which are insoluble in heptane are solubilized therein by a number of detergents.[2]

Some of the most important industrial colloidal systems are nonaqueous. Such are asphalts, lubricating greases, and other petroleum products, coal tar, rubber, paints, lacquers, inks, adhesives, resins and plastics, and many other products. Nevertheless far less quantitative study has been devoted to the nonaqueous colloids than to those in water.

All the phenomena which have become familiar in aqueous colloidal systems have been found in nonaqueous systems with the possible

exception of tactoids and Schiller layers, although these too are as-
sumed by Voet.[3] In nonaqueous systems we observe light scattering,
abnormally high structural viscosity, thixotropy, abnormally low
osmotic pressure, coagulation by diluents or by incompatible colloids,[4]
coacervation, formation of both jellies and gels, hysteresis, protective
action and peptization, syneresis, diffusion, dilatant sediments from
stable suspensions, and other sediments of abnormally high sedimenta-
tion volume, even electrostriction in absolute methanol and electro-
deposition of synthetic resins. Some of these are discussed in the later
chapters.

Co-solvency

Nitrocotton is insoluble in ether and also in alcohol at room tem-
perature. However, a mixture of alcohol and ether is one of the best sol-
vents for nitrocellulose. McBain, Harvey, and Smith [5] suggested that
each of the solvents was capable of dissolving a different portion of
the nitrocellulose molecule and the two together therefore resulted
in complete solution. This was termed *co-solvency*, and the two liquids,
co-solvents. Palit and McBain [6] suggested diagrammatically how this
could be brought about in the case of a mixture of benzene and
propylene glycol dissolving sodium stearate even if no colloid were
formed.

Furthermore, the two co-solvents, which may be almost immiscible
with each other, may *blend* into one liquid phase when dissolving a
material, such as soap, that is insoluble in either alone. Soaps cause
similar blending of methyl alcohol and cyclohexane, and of many other
mixtures of a glycol or dihydroxylic solvent with any solvent for hydro-
carbon such as a hydrocarbon, chlorinated hydrocarbon, ketone, an
alcohol, ester, and the like. Many examples of co-solvency have been
adduced by Palit.[7] Blending and co-solvency are often of practical
importance in industrial processes.

McBain, Harvey, and Smith [8] assembled a large number of ex-
amples already known in 1926 where optimum solvents for nitro-
cotton were made from two nonsolvents acting as co-solvents, or from
two solvents, or from a nonsolvent and a solvent. Co-solvency is useful
in making heavy metal soaps freely soluble in such mixtures as pro-
pylene glycol and chloroform. Many detergents can be used to blend
water with organic solvents, such as hydrocarbons or carbon tetra-
chloride.[9] In the case tested it has been shown that co-solvency and
blending are accompanied by association and the formation of colloid.

Methods of Study [10]

The degree of solvent power has been studied by a variety of methods. Most of the methods of examination described in the preceding chapters are applicable to nonaqueous systems. A few may lose their importance, such as those utilizing electrical conductivity in nonionizing solvents (although even here changes in traces of conductivity may be significant), but these are replaced by other methods, such as dielectric constant. This may distinguish between ramifying aggregates producing structural viscosity and other particles kept separate by layers or chains of solvent,[11] especially when the system is investigated while under shear. All the optical methods of examination described in Chapter 6 are directly available; so is ultrafiltration if the system is not too viscous or solid, and the same proviso applies to the use of the ultracentrifuge.

The writer has always been impressed with the importance of association even with high polymeric colloids. McBain, Harvey, and Smith [12] observed that the best solvents for nitrocotton are those giving the lowest relative viscosity at any definite concentration within the range 0.3–0.5%. Shaking a fresh solution lowered the viscosity. Recovering nitrocotton from acetone by various methods gave a product which when redissolved could have a viscosity either unchanged or greater or less than before.[13] The relative viscosity varies with temperature although the polymers themselves cannot change. All this they referred to breaking down and building up structures between neighboring polymeric molecules; in other words, to varying association of polymeric molecules.

Similarly, from the fact that the osmotic pressure of polyvinyl alcohol solutions depends upon the solvent and is increased by peptizing salts and decreased by precipitating salts, Dobry and Boyer [14] conclude that aggregation and disaggregation are involved. Much evidence of association of tobacco mosaic virus, gelatin, and poly-amino acids, quite different types of large molecules, was adduced by Schachman, Ferry, Eirich, and others at the 1948 National Colloid Symposium.

In somewhat more concentrated solutions the viscosity of solutions of cellulose derivatives increases with no less than the eighth power of the concentration c [15] in approximate accordance with Philippoff's empirical equation:

$$\eta_r = \frac{\eta \text{ solution}}{\eta \text{ solvent}} = \left(\frac{1 + [\eta]c}{8}\right)^8$$

Likewise the apparent viscosity η_0 at very low rates of shear is far greater than that at very high rates of shear η_∞, the difference being attributed to effects of structural viscosity, that is, association.

Methyl cellulose in water sets to a jelly (association) on warming and on cooling reverts to a sol. Lieser, in studying solutions of cellulose xanthate found evidence of micellar association.[16]

"*Solvent power*" has been estimated by various authors using a number of methods. The first and most sensitive is the apparent viscosity just discussed. One frequently used, which was introduced by Mardles, is the amount of a diluent which can be added to the solution before the solute or high polymer is precipitated. However, it is already apparent from the discussion of co-solvency that this criterion has to be used with considerable discretion and even a perfect nonsolvent would not obviate this difficulty since it does not take into account the different kinds of chemical grouping with which various solvents unite. Some writers try to deduce from this indication of solvent power the relative solvation of the colloid, but the term is not used consistently by different writers.

Erbring and Wenstöp [17] investigated a series of alcohols determining their precipitating power on colloidal chloroform solutions of polystyrene, rubber, and cellulose acetate. They found this parallel to μ^2/σ where μ is the polarizability and σ the dielectric constant. This relation has also been stressed in the extensive investigations of Moll.[18] From this and the surface tension of the solvent he could predict whether swelling or solution or no interaction would take place. He also deduced values for solvation.

Further indirect methods of obtaining an indication of solvent power depend upon changes in properties with temperature or upon adding diluents or noting swelling or sorption. Miss King [19] found that the best solvents were those that exerted the greatest retardation or protective action upon the rate of crystallization of the solvent from the solution. It should be recalled that sorption of solvent from the vapor phase usually amounts to only a fraction of the weight of the solute.

The Oliensis [20] spot test is a sorption method for distinguishing the colloidal stability of various constituents of asphalt by noting the order in which they are sorbed on a filter paper.

The *osmotic coefficient* is now one of the most usual methods of study of high polymers in solution. Instead of the osmotic pressure being directly proportional to the concentration as in a regular solution, it rises much more rapidly, as shown in Figure 19·1. Wo. Ost-

wald [21] has expressed osmotic pressure by the approximate empirical equation

$$P = ac + bc^n$$

where a and b are constants, c is the concentration and n is at least 2. By extrapolation to zero concentration the weight M of the kinetically independent particles is deduced:

$$\lim_{c \to 0} \frac{P}{c} = \frac{RT}{M}$$

Schulz [22] had much success in attributing to solvation the deviations from direct proportionality to the concentration. In Figure 19·2, $\frac{P}{c}$ is plotted for a series of nitrocelluloses in acetone up to concentration c of about 4%. Although the molecular weights deduced from the limit varied from 20,000 to 100,000 the lines are all parallel.

Meyer has shown that rubber, gutta percha, and cellulose acetate in different solvents behave similarly.
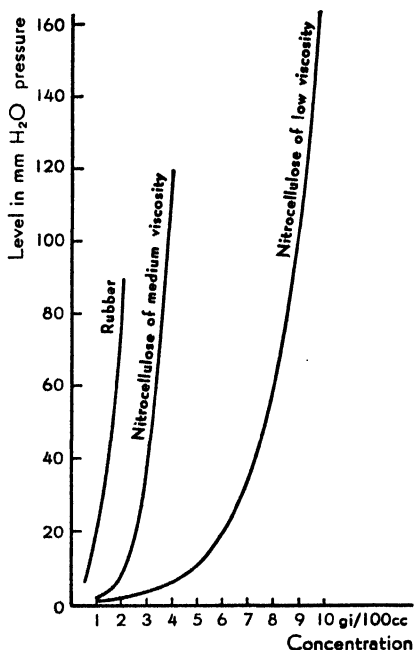
FIGURE 19·1 Rise of osmotic pressure with concentration in a few high polymeric substances. Taken from Mark, H., *Physical Chemistry of High Polymeric Systems*, Text Ref. No. 10.
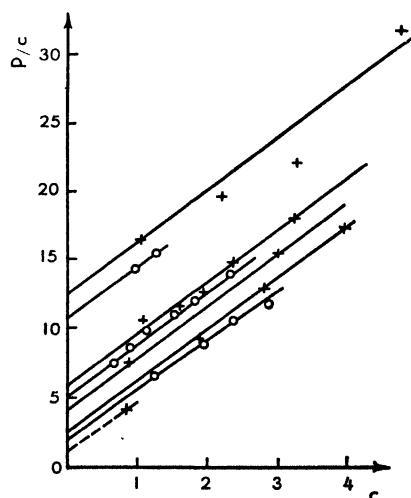
FIGURE 19·2 The function P/c plotted against c for a few nitrocelluloses of different molecular weights. Taken from Mark, H., *Physical Chemistry of High Polymeric Systems*, Text Ref. No. 10.

Sedimentation Volume and Thixotropy

Correns and Winkler [23] found that powdered quartz, feldspar, and fluorspar which, as was pointed out in Chapter 10, are not thixotropic in water, are so thixotropic in benzene that with particles of the right size the system after resting one minute shows no liquid flow when the tube is inverted, though the system is liquid when shaken. They conclude that every substance, if the particles are sufficiently fine, will show thixotropy in a suitable liquid.

Akamatu [24] observed thixotropy with glass, talc, zinc oxide, and calcium carbonate in the nonpolar, nonionizing liquids benzene, toluene, hexane, and carbon tetrachloride. This is evidently due to ramifying aggregates and association of the particles. It goes parallel with a large sedimentation volume, which again indicates loose packing. If the powder is a polar substance, the sedimentation volume in a nonpolar liquid is always larger than that in a polar liquid. On the other hand, with nonpolar graphite it does not depend upon the nature of the liquid. Lampblack in kerosene is thixotropic. Addition of a small amount of polar surface-active substance decreases remarkably both sedimentation volume and thixotropy in a nonpolar solvent. This is attributed by Verwey [25] to increased stabilization of the colloid. The polar additive acts as stabilizing agent. Sediments of intermediate volume have some relation to coacervates; and coacervates in general may well involve a ramifying and scaffolding structure due to contact of particles at points on their surfaces not kept apart by double layers. It appears as if the electrical double layer of the undissociated Helmholtz type in a nonionizing solvent is as important as the combination of Helmholtz and diffuse double layer now accepted for all aqueous systems. It should be pointed out that separation of particles in a sediment merely by a Helmholtz double layer permits close packing and produces dilatancy.

Gallay and Puddington [26] found that finely divided sodium soaps in mineral oil give a sedimentation volume which increases with viscosity and with temperature and decreases with viscosity index, an oleate occupying a greater volume than the stearate. They preferred the conception of a scaffolding structure which accounted for large sedimentation volume and increased apparent viscosity and sometimes a definite yield value, rather than the explanation of Ostwald and Haller [27] assuming lyospheres, that is, particles around which a shell of liquid of varying depth is adsorbed. Both mechanisms may be operative.

Gelation and Coagulation

The formation of gels and jellies is of very frequent occurrence in nonaqueous systems, and it will be discussed more in detail under the heading of soap greases. Examples of nonaqueous jellies are given by rubber in various solvents or by an air-blown petroleum asphalt in light petroleum.

When solutions of asphalt and coal tar pitch are made separately in nonionizing solvents and then mixed, mutual coagulation occurs.[28] Solvent naphtha from distillation of coal tar and mineral spirits from distillation of petroleum are miscible in all proportions and either will dissolve asphalt. However, coal tar pitch dissolves in solvent naphtha but not in mineral spirits. A very small addition of mineral spirits is sufficient to coagulate the coal tar pitch in solvent naphtha. The coagulum is a very hard, brittle substance not like the original pitch. It is possible that the stabilizing agent remains behind in the liquid.

When fresh oil is added to oil which has undergone extreme oxidation in an engine, copious precipitation may result owing to the loss of dispersants and the development of other polar products. Modern detergent-dispersant additives in lubricating oils are usually metal salts of organic acids or of alkyl inorganic acids, highly alkylated aromatic sulfonic acids, and disubstituted dithiophosphoric acids. Other compounds used include highly alkylated phenol sulfides and partially hydrolyzed reaction products of olefins and phosphorus pentasulfide.

Lubricating Soap Greases [29]

Lubrication is often preferably carried out by means of soap greases instead of lubricating oils because the viscosity of the oil changes greatly with temperature, being too thin when hot and too nearly solid when chilled, whereas if the viscosity is structural, due to a mechanical framework of colloidal dimensions, it may be much more nearly unaffected by large changes of temperature. The viscosity of the oil may be altered many-fold without altering the viscosity of the soap grease. Again, soap greases may be easier to apply. Finally, the soap itself may be an excellent lubricant.

A lubricant composed of fat and lime was used as early as 1400 B.C. on wheels of chariots; however, it was not until about 1845 that lime soap was combined with mineral oil to form a grease. Today such

greases are widely used for line shafting, water pumps, plain bearings, and general applications where the temperature does not exceed 70–90° C, above which calcium soap melts and its structure is lost. A calcium soap grease is water resistant and can be made of any desired consistency. It contains a definite, insoluble hydrate of calcium soap. An anhydrous calcium soap would give a very different and less useful product.

Widely used modern soap greases employ sodium soaps in mineral oil. Because of their high melting point they may be used up to much higher temperatures, to 150° or even over 200° C, and they can be made fibrous or hard or soft and smooth as butter. Da Fano [30] in 1929 was the first to prepare alkali soaps in mineral oils and other organic solvents and to modify the properties of the products by means of additives.

Aluminum soaps are available for a limited temperature range, up to about 75° C only, because their texture depends greatly upon the temperature although they are adhesive and insoluble in water. Recently lithium soaps have come into prominence because to the properties of sodium soaps they add more resistance to water. Widely used soap greases often contain mixtures of calcium and sodium soaps. Finally, a few synthetic greases in which neither soap nor petroleum is used will be enumerated below.

The typical behavior of a sodium soap in a nonaqueous solvent whose melting and boiling points permit a sufficient range of observation is as follows: At the highest temperatures the soap dissolves as a simple nonelectrolyte, that is, as ordinary molecules. The best studied example is that of the sodium soaps in pinene,[31] where the molecular weight has been studied by measuring the lowering of vapor pressure. Between 120° and 150° C sodium oleate dissolved in pinene had a mean molecular weight of 306, theory 304. If the temperature is lowered, the solution becomes obviously colloidal and viscous. On further cooling the colloidal sol becomes a clear, transparent jelly. Finally, below the temperature at which the solution becomes saturated, an opaque curd or gel is formed in which the soap has crystallized out in the form of fine fibers enmeshing the system.

Sodium soap greases are gels in which only a trace of soap is dissolved, the rest remaining as insoluble fibers whose degree of comminution and of independence determines whether the grease is buttery or is firm. The presence of these fibers is readily demonstrated by x-ray examination or by the electron microscope after removal of the solvent by extraction.

"Gels" versus "Jellies." Calcium soap greases are also gels in which the oil is merely present in the mass of solid fiber soap. Lithium soap greases are similar. On the other hand, aluminum soap greases are not gels and are never fibrous. They are usually jellies and sometimes liquid crystalline phases. These soaps are good illustrations of the behavior of gels and jellies discussed earlier in Chapter 11. Here again a clear distinction is drawn between transparent jellies which merge insensibly into sols upon change of temperature or concentration and opaque gels in which x-ray examination shows the persistence of the original undissolved solid, and again liquid crystalline phases which are semitransparent, doubly refracting, and plastic. The jellies, except when under strain, are isotropic like the sols; and all of them differ from gels in that the soap has dissolved, so that the original crystals have disappeared and the molecules therein have been reorganized into colloidal particles. A heat of crystallization of several thousand calories is evolved when a gel is being formed, and there is also a less but similar heat effect during the gradual formation of a jelly from a sol. Hence a jelly or a gel cools more slowly and also heats more slowly than the solvent. Formation of acid soap as by addition of 1% oleic acid increases the solubility of the gel fibers so that a nonsyneretic jelly may result. Myristic acid itself in benzene is a jelly. Gels occur where the gel fibers enmesh a solvent, a sol, a jelly, or even a liquid crystal.

The most studied soaps are sodium stearate and sodium palmitate.[32] They have been studied in hydrocarbons, alcohols, fatty acids, glycerols, cresols, terpene, and a number of other solvents. In each of these pure solvents, just as in water, sodium stearate is insoluble at room temperature and scarcely even swells. Blending of the solvents is required to dissolve it. If the temperature is raised the stearate begins to swell and concurrently begins to dissolve. At about 90° it is fully dissolved to form isotropic solutions if dilute, and anisotropic liquid crystalline phase if more concentrated. On being heated to a sufficiently high temperature, the soap becomes miscible with all these solvents as an isotropic phase. The solubility has been discussed by Bondi,[33] and it is rather remarkable that in a number of such cases, such as the self-containing OH or NH_2 groups, the behavior is approximately ideal, just as is the behavior of the soaps in aqueous neat soap phase in spite of their colloidal condition.

The phase diagrams are not reproduced here because there are some unresolved contradictions between the results for highly concentrated systems as studied in different laboratories.

Aluminum dilaurate is an association colloid in benzene. McBain and Working [34] found by osmotic pressure measurements that the number of moles per particle was as follows:

	1%	0.1%	0.01%	0.001%
20° C	573	215	20	6
50° C	681	—	31	10

Therefore the particle weight varies from 300,000 to 2600 with change in concentration. In contrast with the behavior of high polymers the osmotic pressure divided by the concentration decreases rapidly with concentration. The viscosity of these solutions decreases with time, decreases rapidly with dilution, and increases rapidly with rise of temperature (just like methyl cellulose in water). This is further evidence for a structural viscosity due to association.

Some authors have based the interpretation of aluminum soap jellies upon the coordination of aluminum and draw a close analogy with high polymers. This scarcely seems worthwhile when sodium soaps exhibit such similar properties.

Aluminum soap greases are used not merely for lubricating but in some of the most important war weapons as fillers for incendiary bombs and fuels and flame throwers. For these purposes jellies are required. Their properties are greatly modified by additives which can stiffen or liquefy them, retard or accelerate their change in viscosity, and otherwise modify them as desired. The mono-soaps are insoluble and valueless; the tri-soaps do not exist, and the di-soaps, such as aluminum dilaurate $(AlOHL_2)$, are the basis of these jellies. In addition to the dilaurate, dinaphthenates and dioleates are employed in substantial amounts.

Finely divided silica in various forms, and even glass, have the same rigidifying effect on oil as does soap, but unless very finely divided, correspondingly more is required. If the silica particles were fine enough to avoid abrasion, this would constitute the cheapest petroleum lubricating grease. All lubricating greases contain minor additives such as antioxidants, anticorrosive agents, and additives that affect lubrication or viscosity under extreme conditions.

Other Synthetic Lubricants [35]

The extreme conditions encountered by aircraft and in mechanized warfare in different climates have forced the development of a number of lubricating greases not based upon petroleum. Some of these are

based upon polymerized silicones and may contain lithium stearate. Military aircraft employed diesters [36] (di-2-ethylhexyl adipate and sebacate) as a base used between $-68°$ C and perhaps up to $150°$ C. To these and other greases is commonly added some synthetic resin.

A third group of synthetic lubricant hydraulic fluids are designated "UCONS" based upon polyalkylene glycols and derivatives, which may be water soluble or water insoluble and which do not affect natural or synthetic rubber. These have already found very extensive and varied uses.

REFERENCES

Text References

1. LAWRENCE, A. S. C., *Trans. Faraday Soc.*, **34**, 677 (1938)
2. MCBAIN, J. W., MERRILL, R. C. and VINOGRAD, J. R., *J. Am. Chem. Soc.*, **62**, 2880 (1940)
3. VOET, A., *J. Phys. Coll. Chem.*, **51**, 1037 (1947)
4. MCBAIN, J. W. and WORKING, E. B., *J. Phys. Coll. Chem.*, **51**, 974 (1947)
5. MCBAIN, J. W., HARVEY, C. E. and SMITH, L. E., *J. Phys. Chem.*, **30**, 312 (1926); also MCBAIN, J. W., GRANT, E. M. and SMITH, L. E., *ibid.*, **38**, 1217 (1934)
6. PALIT, S. R. and MCBAIN, J. W., *Ind. Eng. Chem.*, **38**, 741 (1946); *J. Am. Oil Chem. Soc.*, **24**, 190 (1947)
7. PALIT, S. R., *J. Am. Chem. Soc.*, **69**, 3120 (1947)
8. MCBAIN, HARVEY, and SMITH, *loc. cit.*
9. PALIT, S. R. and MCBAIN, J. W., *J. Soc. Chem. Ind.*, **66**, 3 (1947)
10. MARK, H., *Physical Chemistry of High Polymeric Systems*, Interscience Publishers, Inc., New York (1940)
    MEYER, K. H., *Natural and Synthetic High Polymers*, Interscience Publishers, Inc., New York (1942)
    FUOSS, R. M., "High Polymers," *Annals of N. Y. Acad. of Sci.*, **44**, 263–443 (1943) "A General Discussion of the Phenomena of Polymerization and Condensation," *Trans. Faraday Soc.*, **32**, 1–412, (1936)
    DOTY, P. M. and MARK, H., *Ind. Eng. Chem.*, **38**, 682 (1946)
11. VOET, *loc. cit.*
12. MCBAIN, HARVEY, and SMITH, *loc. cit.*
13. MCBAIN, J. W. and SCOTT, D. A., *Ind. Eng. Chem.*, **28**, 470 (1936)
14. DOBRY, A. and BOYER, F., *J. Chim. phys.*, **40**, 151 (1943)
15. PFEIFFER, G. H. and OSBORN, R. H., in *Cellulose and Cellulose Derivatives*, Interscience Publishers, Inc., New York (1943), p. 962
16. LIESER, T., *Annalen der Chemie*, **522**, 48, 56 (1936)
17. ERBRING, H. and WENSTÖP, K., *Kolloid-Z.*, **85**, 342 (1938)
18. MOLL, W. L. H., *Kolloid-Beihefte*, **49**, 1 (1939)
19. KING, A. M., *Trans. Faraday Soc.*, **24**, 457 (1928)
20. OLIENSIS, G. F., *Proc. Am. Soc. Testing Materials*, Preprint No. 93 (1941)
21. OSTWALD, WO., *Kolloid-Z.*, **49**, 60 (1929); **51**, 273, 361 (1930); **67**, 211 (1934); *Z. physik. Chem.*, **159**, 375 (1932); also HALLER, W., *Kolloid-Z.*, **56**, 257 (1931)
22. SCHULZ, G. V., *Z. physik. Chem.*, **158**, 237 (1932); **159**, 347 (1932); **160**, 409 (1932); **176**, 317 (1936)
23. CORRENS, C. W. and WINKLER, H. G. F., *Naturwissen.*, **26**, 517 (1938)
24. AKAMATU, H., *Bull. Chem. Soc. Japan*, **13**, 456 (1938)

25. VERWEY, E. J. W., *Rec. trav. chim.*, **60**, 618 (1941)

26. GALLAY, W. and PUDDINGTON, I. E., *Can. J. Res.*, **22**, 16 (1944)

27. OSTWALD, WO. and HALLER, W., *Kolloidchem. Beihefte*, **29**, 354 (1929)

28. BRUCE, H. D., private communication

29. KLEMGARD, E. N., *Lubricating Greases: Their Manufacture and Use*, Reinhold Publishing Corp., New York (1937)

    ELLIOTT, S. B., *The Alkaline Earth and Heavy Metal Soaps*, A. C. S. Monograph # 103, Reinhold Publishing Corp., New York (1946)

    JENČIČ, S., *Kolloid-Z.*, **42**, 69, 168 (1927); **60**, 59 (1932)

    MCBAIN, J. W. and MCCLATCHIE, W., *J. Phys. Chem.*, **36**, 2567 (1932)

    LAWRENCE, A. S. C., *Trans. Faraday Soc.*, **34**, 660 (1938)

    SMITH, G. H., *J. Am. Oil Chem. Soc.*, **24**, 353 (1947)

30. DA FANO, E., *Giorn. chim. ind. applicata*, **11**, 199 (1929); *Ind. olii min e grassi*, **9**, 105 (1929)

31. PRASAD, M., HATTIANGDI, G. S. and VISHVANATH, C. V., *Proc. Indian Acad. Sci.*, **21**, 64 (1945)

32. VOLD, R. D., LEGGETT, C. W. and MCBAIN, J. W., *J. Phys. Chem.*, **44**, 1058 (1940)

    LEGGETT, C. W., VOLD, R. D. and MCBAIN, J. W., *J. Phys. Chem.*, **46**, 429 (1942)

    GALLAY, W. and PUDDINGTON, I. E., *Can. J. Res.*, **21B**, 202, 211 (1943)

    PRASAD, M., and HATTIANGDI, G. S., *Proc. Indian Acad. Sci.*, **21**, No. 1, A. (1945)

    DOSCHER, T. M. and VOLD, R. D., *J. Coll. Sci.*, **1**, 299 (1946)

33. BONDI, A., *J. Chem. Phys.*, **16**, 157 (1948)

34. MCBAIN and WORKING, *loc. cit.* See also SHEFFER, H., *Can. J. Res.*, **26B**, 481 (1948)

35. *Symposium on Synthetic Lubricants*, 50th Ann. Meeting Am. Soc. Test. Materials, June 16, 1947, Special Tech. Bull. # 77

36. ATKINS, D. C., BAKER, H. R., MURPHY, C. M. and ZISMAN, W. A., *Ind. Eng. Chem.*, **39**, 491 (1947)

X–RAY AND OTHER METHODS OF STUDYING

COLLOIDS AND HIGH POLYMERS

Association is the fundamental phenomenon in colloid behavior. A perfect gas would be the only system in which associating behavior is wholly absent. In a real gas there is a tendency for evanescent clustering or temporarily inelastic collisions of molecules. The common conception of a so-called normal liquid is one in which the molecules, although close-packed, retain complete freedom of rotation and movement. However, the phenomenon of cybotaxis, referred to in Chapter 4, shows a temporary association or orientation of neighboring molecules, and in some liquids this tendency has long since been recognized. They are called associated polar liquids, such as the fatty acids and water. Even in the vapor state a fatty acid such as acetic acid tends to associate so that a pair of individual, single molecules or monolayers become a dimer held together by hydrogen bonds as in the formula:

$$\text{R—C} \underset{O \rightarrow H—O}{\overset{O—H \leftarrow O}{<>}} \text{C—R}$$
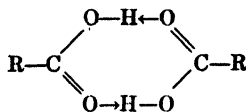
This whole book has dealt with cases where association has produced polymolecular *units* or colloidal micelles, both where these units are kept separate, as in sols, and where they are more or less linked, as in gels and jellies. A still more extensive organization is that of some gels and glasses. Next come fibers, such as the crystalline fiber silicates, and also the analogous structures of drawn wires, of muscle, of other proteins, and of cellulose derivatives. These latter are called high polymers, or macromolecules, because their individual molecules are

a result of the linking together by actual chemical bonds of a large number of ordinary chemical molecules to form a single large molecule, the molecule of the high polymers, which then shows associative effects. A single high polymer may be made up of one or more kinds of monomers (Chap. 25).

Other very numerous groups of systems whose properties are conditioned by association and cybotaxis over large areas are the anisotropic liquids,[1] the liquid crystals or crystal liquids, or mesomorphic phases, as they have been variously termed, in which the unit itself is often a highly organized colloidal particle. It is beyond the scope of this volume to describe these fascinating materials in great detail, but they embrace thousands of pure organic compounds and the varieties of possible structures are surprisingly numerous.

Lastly, we may mention crystals which may be made of atoms, as illustrated by the metals and diamond, or the lamellar structure of graphite, or the spiral structure of plastic sulfur.[2] Finally, there are kinds of crystals in which the units are molecules.

Many methods of examination of colloids and high polymers have already been discussed in the previous chapters. However, many more are available, and significant information is drawn from such fields as synthetic organic chemistry, mechanical tests, examination of refractive index, dielectric constant, dipole moment, solubility and volatility, magnetic susceptibility,[3] and, last but not least, x-ray and electron diffraction.

Although x-ray crystallography [4] is a field for specialists, every student of colloids should know of Bragg's law and should have an acquaintance with the general appearance and significance of powder photographs, fiber photographs, single-crystal rotation photographs, and Laue photographs.[5] It would also be of great advantage for him to familiarize himself with the characteristics of the six crystallographic systems: isometric (cubic), tetragonal, hexagonal, orthorhombic, monoclinic, and triclinic, and the significance of the Miller indices.

FIGURE 20·1 Diagram illustrating Bragg spacings in an hexagonal array. Taken from Astbury, *Fundamentals of Fibre Structure*, Text Ref. No. 6.

A most readable nontechnical introduction to the application of x-ray examination to crystals and high polymers is that of W. T. Ast-

bury,[6] from which Figure 20·1 is taken to illustrate some "Bragg spacings" in an hexagonal array. Figure 20·2 shows how these Bragg spacings can result in diffracted x rays which go to the photographic plate, ionization chamber, or Geiger counter, at an angle of 2θ to the incident

∴ Path difference = 2d sin θ
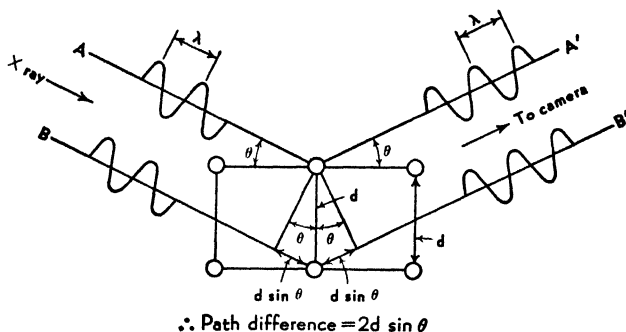
FIGURE 20·2 Diagram showing how Bragg spacings can result in diffracted x rays. d = distance between scattering centers; λ = wave length of incident radiation (x rays); path difference between wave AA′ and BB′ = 2d sin θ. If path difference = integral multiple of wave length (nλ) the waves are reinforced. Otherwise they will be out of phase and will interfere. Therefore path difference = nλ = 2d sin θ for reinforcement. Bragg's law: nλ = 2d sin θ.

x-ray beam. The x-ray beam is made as parallel as possible by being collimated through two narrow slits or pinholes. Most of the x rays pass right through the specimen, and to avoid undue darkening of the photographic plate this main beam is intercepted and absorbed by a small button of lead, which therefore casts a white shadow on the middle of the photograph. The appearance of a Laue photograph of a single fixed crystal is illustrated in Figure 20·3. Each spot is the result of a Bragg spacing in the three-dimensional crystal lattice.

From the Bragg equation illustrated in Figure 20·2

$$n\lambda = 2d \sin \theta$$

it will be noted that the longer the Bragg spacing, that is, the greater the distance d between the parallel planes in which scattered centers of atoms lie, the less the angle of diffraction. Thus with very long spacings the reflection corresponding to the first order of the long spacing may lie rather close to the blanked out center. However, its higher orders, where different integral values are given to n, will lie farther out just in those proportions. Short Bragg spacings will be diffracted at high angles.

In the Laue photograph, x rays of all wave lengths are present in the beam. It is much more usual to employ monochromatic radiation obtained by using say a filter of nickel foil or by reflection from another
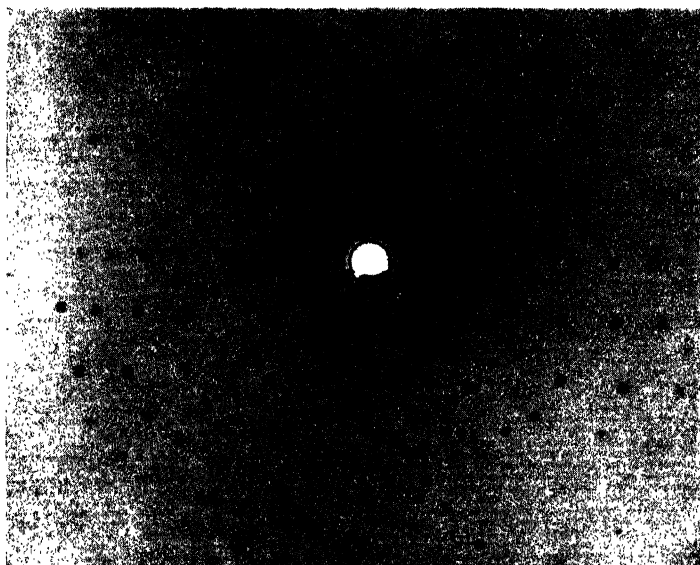
FIGURE 20·3 Typical Laue photograph of a single fixed crystal.
Courtesy of W. T. Astbury.
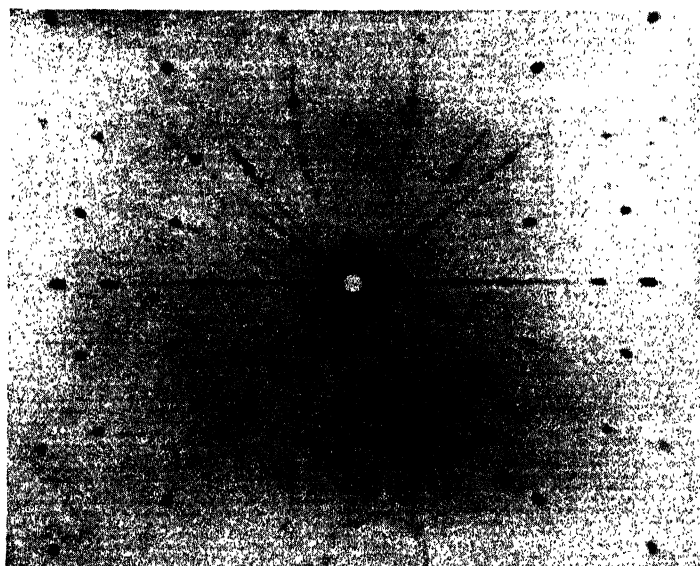
FIGURE 20·4 Typical single crystal rotation photograph.
Courtesy of W. T. Astbury.

crystal. In this case a single crystal has to be rotated about an axis, usually vertical, in order that all planes may have an opportunity of being brought to the exact angle for Bragg reflection. Such a single crystal rotation photograph is illustrated in Figure 20·4, and again its characteristics should be noted. All of the spots lie on a series of hyperbolas, symmetrical about the vertical axis and the horizontal equator. With a different arrangement of the camera these hyperbolas all become straight lines and the distance between these horizontal lines or layer lines corresponds then to Bragg spacings in the vertical direction. This is important, as will be shown, because a fiber photograph has similar characteristics and these layer lines give the spacing along the axis of the fiber, whereas the spots along the equator measure spacings at right angles to the fiber.

Now if a single crystal were to be rotated into all possible orientations, each of the spots on the Laue or fiber photograph would be drawn out into a complete circle because for each the angle of deviation 2θ from the main beam would be constant. The same purpose is achieved in practice by using a fine powder of crystals so that all orientations will be represented. Such a powder photograph of hexanolamine myristate taken with monochromatic radiation is illustrated in Figure 20·5. Although powder photographs are of great value as unique fingerprints of the given substance, which can lead to its comparison and identification, powder photographs suffer from the great disadvantage that in general the crystal characteristics cannot be unambiguously determined because of the impossibility of distinguishing between the lines representing spacings in the different crystal directions.

FIGURE 20·5 Typical powder photograph. Hexanolamine myristate, monochromatic radiation, from the Stanford laboratories of the author.

A rotating single crystal photograph or the very similar photograph that is obtained by x raying a vertical bundle of parallel crystalline fibers is much more informative, because, as already pointed out, the distance between the layer lines identifies spacings in the direction of the fiber axis. If such a bundle of fiber crystals, whether of asbestos or a fiber such as cellulose, stretched protein, or stretched rubber, is pho-
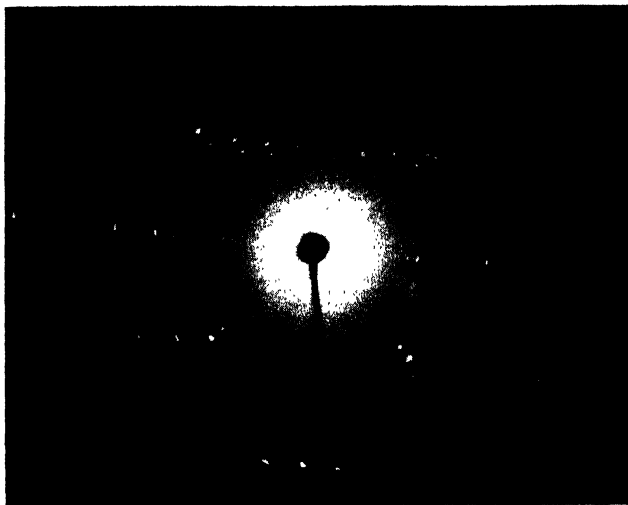
FIGURE 20·6 Typical fiber photograph. Hexanolamine caprylate, from the Stanford laboratories of the author.

tographed, it is of course unnecessary to rotate the fibers in order to get a photograph similar to Figure 20·4. Such a fiber photograph is reproduced in Figure 20·6, which is a photograph of the amine soap, hexanolamine caprylate. Similar fiber photographs will be presented in the ensuing chapters on natural and synthetic high polymers. The soap fibers present the advantage that the molecules are so long in comparison with their breadth and thickness that the small-angle diffractions lying on the equator of such a photograph as Figure 20·6 can be identified as the long spacings. We therefore have distinguished the long or "c" spacings from an inspection of Figure 20·6 which, in this case, must be at right angles to the axis of the fiber. In the next chapter it will be emphasized that fibers of high polymeric molecules have the molecular axis, not like soap, but parallel to the direction of the fiber.

We have also identified the "b" spacings in the direction of the fiber axis, leaving the spacing in the third or "a" direction to be deduced by a process of elimination. It is therefore seen that one can obtain considerable detailed information with regard to crystal characteristics even when single crystals are not available, as is unfortunately the case with high polymers. Finally, the x-ray beam may be made so fine that the center of a single fiber such as nylon and its outside layers can be studied independently [7] or so long that a spacing as 640 Å in collagen can be observed in 33 orders. [8]

The most important pioneer work was carried out with cellulose fibers, as illustrated in Figure 20·7. We recall that sharpness of spots and lines corresponded to perfection of spacing and orientation. Likewise it must be mentioned that the spacing must repeat a certain number of times before the lines become sharp. Indeed, from the diffuseness of lines from good crystals one can infer the size of the crystals if they are sufficiently small; however, a *particle* may be made up of many such crystallites. Kratky [9] and Guinier [10] have developed a method of studying *particle* size from low angle scattering in dilute suspensions by applying to x rays the laws governing the production of halos (such as those around the moon in thin high clouds) with ordinary light. It may be applied to demonstrate processes of subdivision and increase of surface in the activation of catalysts. This important technique has been applied to the study of the magnitude and arrangement and interpretation of periodicities observed in natural and synthetic fibers and plastics.[11]

FIGURE 20·7 Fiber photograph of natural ramie. Taken from Astbury, *Fundamentals of Fibre Structure*, Text Ref. No. 6.

Now Figure 20·7 shows blurred lines which are drawn out appreciably into arcs and which are elongated more in a horizontal than in a vertical direction. Hence it can be concluded that the crystals of which the textile fibers of Figure 20·7 were composed are much longer than they are thick and that they lie lengthwise along the fiber. The arcs show that the crystals are not all strictly parallel to the fiber axis but may be inclined at small angles. In the next chapter we shall see that cellulose in the form of cotton has a spiral structure which can be deduced from the lengths of the arcs into which the spots are drawn.

X-ray analysis has therefore shown that natural vegetable fibers are comprised of crystals of cellulose and that these are very long and very thin and parallel to the fiber axis. The significance of the repeating distance along the fiber axis, 10.3 Å, is discussed in the next chapter.

X-ray examination, in favorable cases, can therefore not only detect crystalline matter and "fingerprint" it, but also determine the

degree of crystallinity of a particular substance.[12] It can also determine the actual positions of molecules and atoms and their dimensions.

It is found that, in general, atoms have fixed radii, that they may be polarized, and that polarization distorts the angles between adjacent atoms more than it affects their distance apart.[13] The primary valences of carbon and silicon are at the corner of tetrahedra and therefore have a bond angle of 109.5°. This accounts for the zigzag chain of hydrocarbons and their derivatives, for the spirals of silicon and oxygen atoms in quartz, or for layers in silicates and clay.

Atoms connected by primary valences have a distance apart from center to center of between 1 and 1.6 Å. For example, the normal distance apart of carbon atoms united by a primary valence is 1.54 Å, with a double bond 1.34 Å, as in ethylene, and with a triple bond 1.20 Å, as in acetylene. The primary C—N bond is 1.47 Å and of N—N 1.47 Å. Diamond has therefore an interatomic distance of 1.541 Å. The spacing between hydrogen and carbon, nitrogen and oxygen, and oxygen and hydrogen is 1.09, 1.01, and 0.97, respectively. That between oxygen and carbon is 1.42 Å.

Benzene maintains a "resonating" compromise between single and double bonds giving the distance between carbon atoms of 1.39 Å. Graphite, like naphthalene and anthracene, has a C—C spacing in the graphite layers of 1.41 Å, which is two thirds of the way between a double and single bond. However, it should be noted that the distance between successive lamellae is 3.4 Å. A hydrogen bond is 2.5 Å and a so-called hydroxyl bond, as in ice, is 2.76 Å. Molecules that are not bound by chemical valences keep at least 3.4–4 Å apart. Such spacings are therefore characteristic of purely physical forces such as those of van der Waals which are operative in cohesion, adhesion, and the physical type of adsorption. Hence the Bragg spacings take on direct chemical significance, and *vice versa*.

Bearing these distances in mind, one sees that the densities of different types of matter can be approximately predicted. Thus the densities of an ordinary molecular compound will lie between 0.6 and 1.5; that of a lamellar crystal, as in graphite, 2.2; that of a chain or fiber, 1.5. Similarly these bond distances and chemical bonds go parallel with the bond energies, primary valences corresponding in general to between 60 and 181 kilogram calories as can be determined from heats of combustion and from band spectra. The strongest co-valence,[14] which is that of diamond, is 135 kilogram calories per mole. An electro valence, as in an ionic crystal of sodium chloride, is 181 kilogram calories per mole; metal lattices as in sodium, 26; permanent dipoles, 7.

Lastly, secondary valences between chemically independent molecules, as for example in liquids, have a bond energy of only 1–15 kilogram calories, which can be checked by determining heats of vaporization, sublimation, and sorption. Tensile strength will be discussed later.

As another deduction from such data it is clear that swelling of various substances cannot take place in directions that are bound by primary valences but can only take place where the distance already is at least 3.4 Å, as in graphite and graphitic acid, the chemical spacing remaining completely unaltered.[15] Similarly sorption by charcoal does not alter its primary C—C bonds.[16]

The Debye-Scherrer powder x-ray method was first applied to a colloid in 1920, but the specimen was the dried residue of colloidal gold. This proved that the interior of the colloidal particles consisted of true metallic crystals of gold. Using a similar method, Freundlich [17] studied the constitution of dispersed particles and found that those of colloidal gold and silver were small crystals of the metals; Bredig's copper sols were CuO; and freshly prepared sols of alumina, zirconia, thoria, sulfur, and arsenic trisulfide had amorphous particles. Graham's ferric hydroxide was a basic ferric chloride, and a ferric hydroxide sol made by oxidation of iron carbonyl was $FeO(OH)$. This, of course, referred solely to the interior of the particles and not to their surface or their stabilizing agents. Freshly precipitated silica gel contains crystalline centers of cristobalite type of colloidal dimensions.

It was not until 1936 that actual colloidal sols were directly examined by Milligan and Weiser.[18] They found that the common inorganic hydrosols consist of crystalline particles (apart from stabilizing agent) such as the following:

Alumina	γ-$Al_2O_3 \cdot H_2O$
Ferric oxide	α-Fe_2O_3 or β-$FeOOH$
Stannic oxide	SnO_2
Protected silver	Metallic silver
Indium hydroxide	$In(OH)_3$
Negative silver iodide	AgI
Titanium dioxide	TiO_2
Yellow and red arsenic sulfide	As_2S_3

Precipitated gels of ferrocyanides of copper, manganese, cobalt, and nickel consist of the normal salts, whereas those of silver, zinc, and cadmium exhibit additional diffraction lines.[19] Precipitated hydrous gallium oxide may be amorphous but may transform to a definite crystalline monohydrate.[20]

REFERENCES

Text References

1. *Trans. Farad. Soc.*, **29**, 881–1085 (1933) and *Z. Krist.*, **79**, 1–347 (1931)
2. MEYER, K. H. and GO, Y., *Kolloid-Z.*, **70**, 19 (1935)
3. BRINDLEY, G. W. and HOARE, F. E., *Nature*, **135**, 473 (1935)
4. WYCKOFF, RALPH W. G., *The Structure of Crystals*, 2nd ed., Chemical Catalog Co., New York (1931)
   DAVEY, W. P., *A Study of Crystal Structure and Its Applications*, McGraw-Hill Book Co., New York (1934)
   RANDALL, J. T., *The Diffraction of X-Rays and Electrons by Amorphous Solids, Liquids and Gases*, Chapman & Hall, London (1934)
   BRAGG, W. L., *Atomic Structure of Minerals*, Cornell University Press, Ithaca, N.Y. (1937)
   STILLWELL, CHARLES W., *Crystal Chemistry*, McGraw-Hill Book Co., New York (1938)
   CLARK, GEORGE L., *Applied X-Rays*, McGraw-Hill Book Co., New York (1940). Special attention is given to natural and synthetic high polymeric materials.
   BUERGER, M. J., *X-Ray Crystallography*, John Wiley & Sons, Inc., New York (1942)
   WAHLSTROM, E. E., *Optical Crystallography*, John Wiley & Sons, Inc., New York (1943)
   BUNN, C. W., *Chemical Crystallography*, Oxford University Press, London (1945)
   ZACHARIASEN, WM. H., *Theory of X-Ray Diffraction in Crystals*, John Wiley & Sons, Inc., New York (1945)
   SPROULL, WAYNE T., *X-Rays in Practice*, McGraw-Hill Book Co., New York (1946)
   COMPTON, A. H. and ALLISON, S. K., *X-Rays in Theory and Experiment*, D. Van Nostrand Co., New York (1935)
5. STILLWELL, C. W., *J. Chem. Ed.*, **10**, 590–667 (1933); **11**, 159 (1934), for example
6. ASTBURY, W. T., *Fundamentals of Fibre Structure*, Oxford University Press, London (1933)
7. FANKUCHEN, I. and MARK, H., *J. Appl. Phys.*, **15**, 364 (1944)
8. BOLDUAN, O.E.A., 1948
9. KRATKY, O., *Naturwissen.*, **26**, 94 (1938); *Kolloid-Z.*, **96**, 301 (1941)
10. GUINIER, A., *J. chim. phys.*, **40**, 133 (1943); *Ann. phys.*, **12**, 161 (1931)
11. FANKUCHEN and MARK, *loc. cit.*
12. MARSDEN, S. S., MYSELS, K. J. and SMITH, G. H., *J. Coll. Sci.*, **2**, 265 (1947); MYSELS, K. J. and MCBAIN, J. W., *J. Phys. Coll. Chemistry*, **52**, 1471 (1948); MARSDEN, S. S., MYSELS, K. J., SMITH, G. H. and ROSS, S., *J. Am. Oil Chem. Soc.*, **25**, 454 (1948)
13. For a detailed discussion see PAULING, L., *The Nature of the Chemical Bond*, Cornell University Press, Ithaca, N.Y. (1939)
14. UBBELOHDE, A. R., *Nature*, **145**, 488 (1940)
15. HOFMANN, U., FRENZEL, A., WILM, D. and CSALÁN, E., *Kolloid-Z.*, **61**, 297 (1932)
16. BERL, E. and REINHARDT, L., *Z. physikal. Chem.*, **166A**, 81 (1933)
17. FREUNDLICH, H., *Ber.*, **61B**, 2219 (1928)
18. MILLIGAN, W. O. and WEISER, H. B., *J. Phys. Chem.*, **40**, 1095 (1936)
19. WEISER, H. B., MILLIGAN, W. O. and BATES, J. B., *J. Phys. Chem.*, **42**, 945 (1938)
20. MILLIGAN, W. O. and WEISER, H. B., *J. Am. Chem. Soc.*, **59**, 1670 (1937)
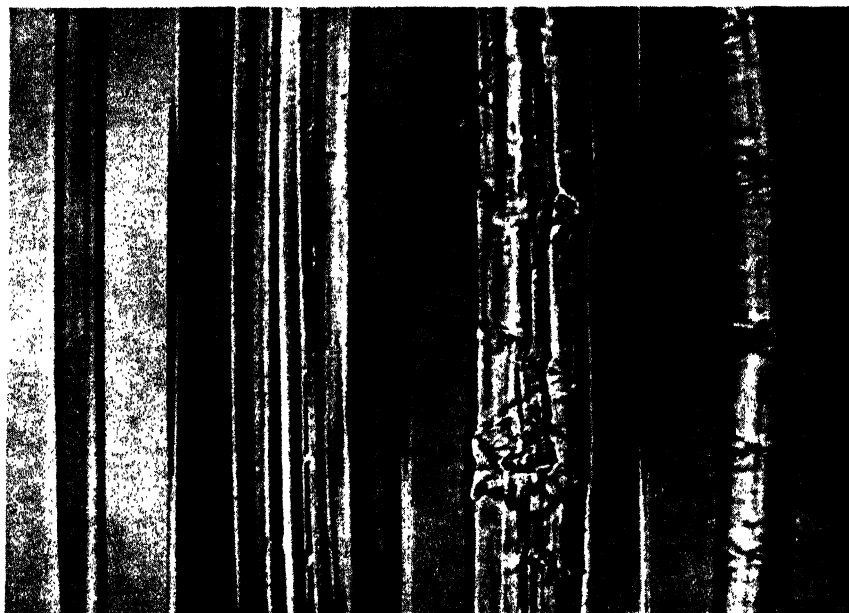
Natural silk   Viscose   Ramie   Flax
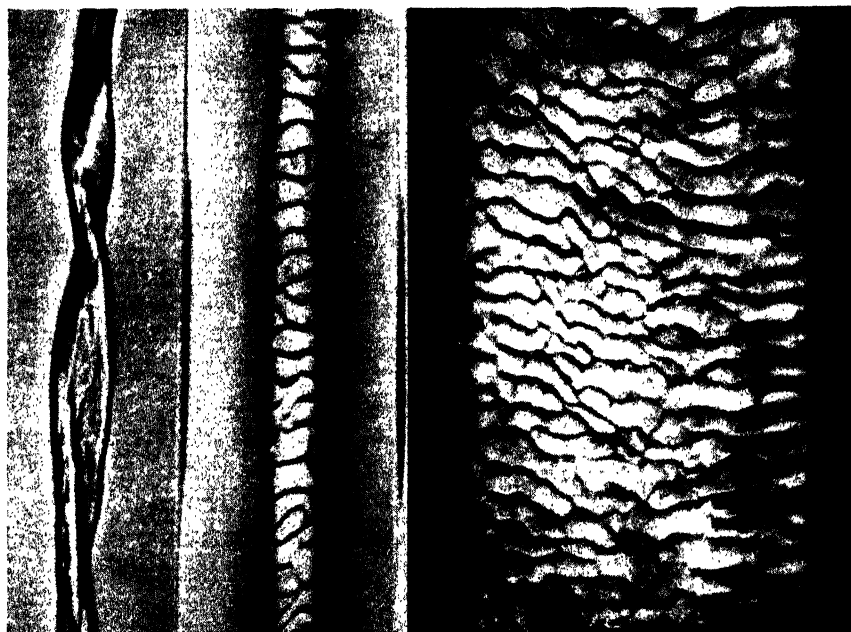
Cotton   Merino wool   Human hair

FIGURE 21·1 Typical textile fibers as seen under the microscope. Taken from the frontispiece of Astbury, *Fundamentals of Fibre Structure*, Chap. 20, Text Ref. No. 6.

CELLULOSE AND ITS DERIVATIVES

Typical textile fibers are illustrated in Figure 21·1. Of these, ramie, flax, cotton, and viscose consist of cellulose, like jute, sisal, hemp, wood, a cell wall of valonia, etc., which are largely cellulose.

It is difficult to make definitive progress in the study of colloidal material until its chemical nature and structure have been determined. Cellulose was the first of the high polymers to have its chemical and x-ray structure elucidated. (See Fig. 20·7.) Sponsler and Dore [1] combined the chemical results of Haworth with their own pioneer x-ray examinations to fix the structure of native cellulose. Like Herzog [2] they found the structural repeating unit to be C_{12}, but unfortunately they interchanged the O with the CCH_2OH group in one of their glucose residues. Hibbert and Barsha [3] found that acetobacter xylinum can transform glucose into genuine cellulose.

Cellulose is made up of glucose molecules condensed in pairs to form cellobiose molecules, as shown in the illustration in Figure 21·2, and the cellobiose units themselves are again condensed into extremely long molecules of cellulose.

In the previous chapter (Fig. 20·7) is given the x-ray fiber photograph of ramie from which it could be deduced that the crystallites were long, thin, and parallel to the direction of axis of the fiber. Now the significance of the repeating unit of 10.3 Å appears. It is that of the cellobiose, and it is at once noticed that there is no spacing corresponding to the actual length of the cellulose molecule but only to the repeating patterns in that molecule. With all forms of cellulose, the principal framework of plants, the repeating pattern is that of cellobiose. The question immediately arises as to the length and molecular

weight of the cellulose molecule. In other words, what is the degree of polymerization (often termed D.P.) or the number of units of cellobiose in one cellulose molecule?

There are six common methods for determining degree of polymerization. All of these methods, with the exception of chemical analysis and end groups, have been discussed in previous chapters. Examination of Figure 21·2 shows that on each glucose residue of cellobiose or cellulose there are three hydroxyl groups available for chemical reaction, for example, substitution by methyl groups or conversion into nitrate or acetate, etc. Now the termination of the molecule at each end provides one more hydroxyl group, and therefore an exhaustive methylation of cellulose with the resulting proportion of tetramethyl derivatives will give the number of end groups. In the case of cellulose this is simple, because the molecule is linear and not branched and therefore has only two ends. In starch, discussed in Chapter 22, the branched molecules supply many such terminal end groups.

The six methods just referred to are listed in the following tabulation, together with some typical values obtained thereby:

Method and Investigator	D.P.	M.W.
1. *Ultracentrifuge*, using a solvent, although no really good solvent has been found for cellulose:		
Stamm	164	56,000
Kraemer	320–880	110,000–300,000
Svedberg *		
Unbleached cotton linters	9,200	1,500,000
Raw Georgia cotton	6,200	1,000,000
Nettle fiber cellulose	10,800	1,760,000
Ramie	11,300	1,870,000
Sulfite wood pulp	2,900	460,000
2. Osmotic Pressure:		
Buchner	877	30,000
3. End Groups:		
Haworth	877 0.6% tet. Me	30,000
4. X-Rays (size of crystallites):		
Meyer and Mark	877	30,000
5. Viscosity	280	9,576
6. Light Scattering (acetate)		63,000–173,000

* Gralén, N. and Svedberg, T., *Nature*, 152, 625 (1943)

Of these methods, ultracentrifuge and viscosity give a weight average. Osmotic pressure and end groups give a number average. "Ultracentrifuge methods measure only particle size, and therefore depend on the degree of aggregation of the molecules."[4] Gralén and Svedberg also measured viscosities for the samples quoted in the preceding table

Cellulose linkage   Cellobiose
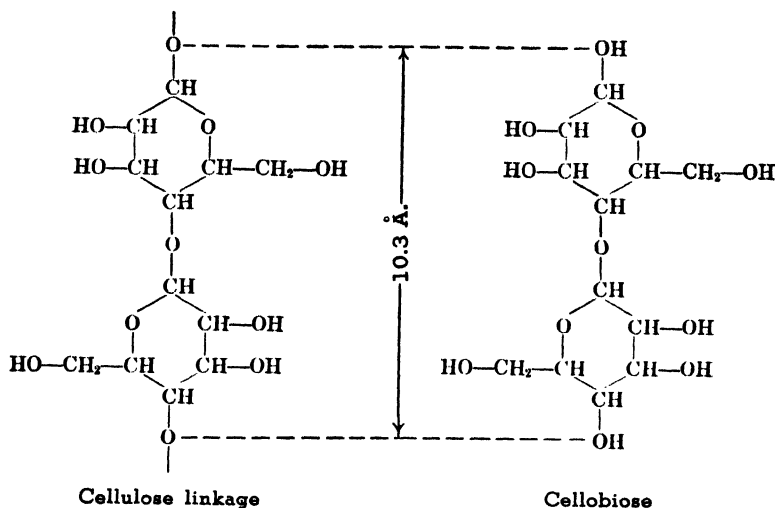
FIGURE 21·2 Cellulose, made up of glucose molecules condensed in pairs to form cellobiose molecules.

and noted discrepancies of many fold in the calculated D.P's. They state:

> It is obvious that our results do not correspond to Staudinger's hypothesis of a linear relation between relative viscosity and molecular weight. At least for these giant molecules, the molecular weight increases very much faster than viscosity.

TABLE 20 *

POLYMERIZATION DEGREES AND MOLECULAR WEIGHTS OF SOME HIGH POLYMERS

Material Investigated	Polymerization Degree	Molecular Weight
Native cellulose in cotton, ramie, or wood	2,000–3,000	300,000–500,000
Cellulose in bleached cotton linters	1,000–1,500	150,000–230,000
Cellulose in purified wood pulp	800–1,200	120,000–200,000
Regenerated cellulose in rayon	500– 600	75,000–100,000
Regenerated cellulose in staple fiber	.400– 500	60,000– 75,000
Regenerated cellulose in cellophane	300– 400	50,000– 60,000
Native rubber in Hevea latex	2,000–3,000	140,000–210,000
Rubber after being milled in air	800–1,000	55,000– 70,000
Nitrocellulose used for molding	1,500–2,500	400,000–700,000
Nitrocellulose used for extrusion	600–1,200	150,000–300,000
Nitrocellulose used for coatings	200– 400	50,000–100,000
Polystyrene for plate casting	2,500–4,000	250,000–400,000
Polystyrene for injection molding	1,200–1,800	120,000–180,000
Polystyrene for coatings	800–1,200	80,000–120,000
Polyvinyl chloride (koroseal)	3,000	250,000
Polyisobutylene (vistanex)	2,000–3,000	120,000–200,000
Polyhexa methylene adipamate (nylon)	150– 300	16,000– 32,000

* Mark, H., *Amer. Scientist*, **31**, 104 (1943)

From the sedimentation data it can also be deduced that the longer the molecules are the more do they tend to become curled.

It is of interest to compare these values with some others for high polymers, as is done in Table 20.
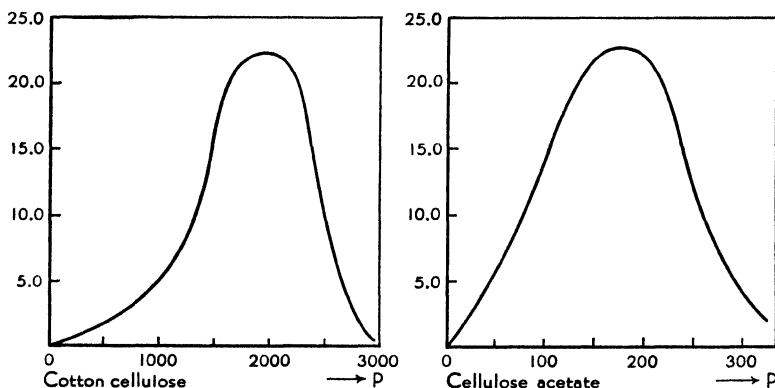
FIGURE 21·3 Diagram showing experimental differential distribution curve of cotton cellulose and cellulose acetate, respectively.

Cellulose is polydisperse, as is illustrated in Figure 21·3.

Now, in more detail, the x-ray structure of cellulose: The unit cell dimensions of cellulose and of cellulose hydrate — that is, mercerized cellulose as precipitated from solution or liberated from its compounds with alkali — are given in the following tabulation. Cellulose III is also included. It should be remembered that cellulose hydrate has the same chemical composition as native cellulose and is free of water. Its x-ray diagram is given in Figure 21·4.

| | Direction | | | Monoclinic Angle |
	a	b	c	
Native cellulose	8.35 Å	10.3 Å	7.95 Å	84°
Mercerized cellulose II	8.14	10.3	9.14	62
Cellulose III	7.48	10.3	8.61	58
High temperature cellulose IV	8.14	10.3	7.90	90

From an examination of the tabulation, the first point to note is that b is the direction of the fiber axis, which is comprised of the cellobiose residues, length 10.3 Å, linked in a long chain by primary valences. This linkage and spacing is not altered by chemical reactions, by dyeing, or by other modifications of cellulose. Such reactions are called topochemical reactions. The tensile strength in the direction of the

axis is about 20 kg/sq mm, but very special portions of fiber more nearly approach the theoretical limit for disrupting primary valences of 800 kg/sq mm.

It may be of historical interest that Bragg [5] observed that the Sponsler and Dore unit cell, although differently described, was mathematically identical with that of Meyer and Mark,[6] which is given in Figure 21·5. The bond energies in the three directions are indicated in calories.

The arrangement of cellulose chains relative to each other as viewed from the direction of the "x" axis is shown in Figure 21·6.[7] Considering the native cellulose, one sees that the parallel molecules are connected in the a direction by hydrogen bonds but in the c direction merely by van der Waals forces. It is therefore at once apparent why cellulose fibers cannot swell in the b direction, but little in the a direction, and that swelling must occur in the c direction. Wood swells laterally but not lengthwise. Ramie swells in liquid ammonia [8] perpendicular to the fiber axis. Removal of the ammonia by water can restore either the original cellulose or a hydrate cellulose structure as desired.

Cellophane does not swell lengthwise, only about 12% sideways, but may increase in thickness by several fold.[9] Electron diffraction shows the Cellophane film to be partly amorphous but crystalline in patches, the surface of the film being the 101 plane.
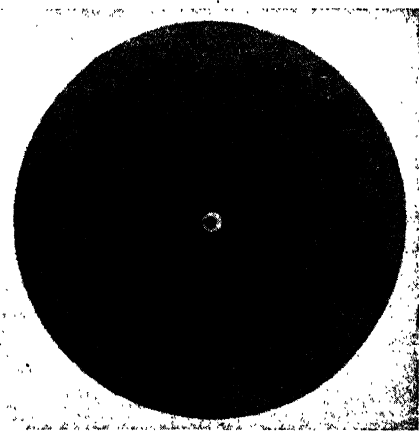
FIGURE 21·4 X-ray fiber photograph of cellulose rayon (mercerized cellulose). Courtesy of W. T. Astbury.

FIGURE 21·5 Representation of the monoclinic elementary cell of cellulose. Three chains of glucose units are shown, running parallel to the b axis (fiber axis). The bond energies in the three directions are indicated in calories. After Mark, H., Chem. Rev., 26, 182 (1940).
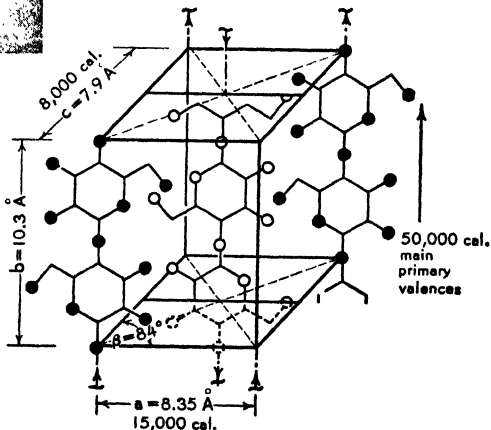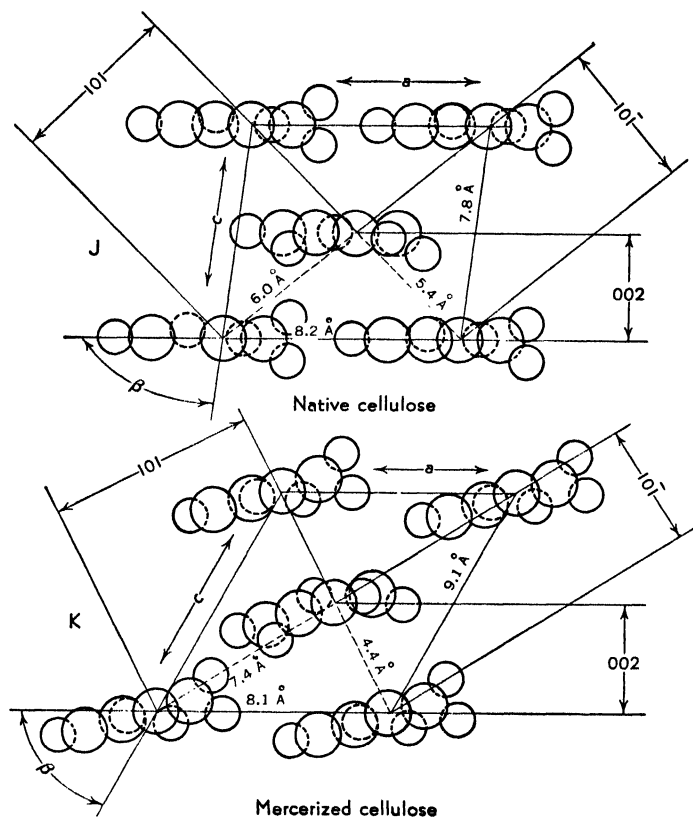
FIGURE 21·6 View down b axis of unit cell of native and mercerized cellulose showing the location of a and c axes, angle beta, and the 101 and 002 planes. Taken from Ott, E., *Cellulose and Cellulose Derivatives* (see References).

It is generally true that if fibers of high polymers are arranged in parallel, their tensile strength and their swelling behavior will obey patterns similar to this.

Figure 21·7 shows in greater detail the arrangement of the glucose residues in the fiber axis, depicting the way in which they are normally somewhat contracted by hydrogen bonds.[10] It should be noted that adjacent cellulose molecules are directed in opposite ways.

The x-ray patterns of cellulose are of different degrees of diffuseness, indicating that the crystallinity of cellulose, like that of other high polymers, is by no means perfect. This has been interpreted in various ways. The earliest interpretation was that cellulose was built up of separate crystallites which were unfortunately termed micelles, leading to confusion with the use of the term throughout this book.

Such a structure is illustrated in Figure 21·8A. This view has been abandoned. The second opposite extreme was that of continuous structure, as shown in Figure 21·8B. This again does not account for the fact that cellulose and other high polymers contain large amounts of amorphous matter. The presently accepted hypothesis is that of Figures 21·8C and D, known as the network and fringe structure, or as the fringe micellar structure. The oriented portions of cellulose have been estimated as approximately 50 × 50 × 600 Å each.

Fankuchen and Mark state:[11] "The average diameter of the micelles in this material (ramie) is between 70 and 100 Å, while their average length exceeds 600 Å." Between these lies amorphous cellulose with its greater power of sorption and swelling. Hermans and Weidinger [12] point out that this sharp differentiation between crystalline and amorphous is "but a rough approximation." They show how different components of order in one direction may be successively disturbed without necessarily affecting order in the other directions. Some such explanation may also apply to the progressive variability and inhomogeneity of aluminum dilàurate.[13]

Kratky [14] has shown by small angle x-ray scattering that rolled out hydrate cellulose is comprised of leaflike cells about 60 Å thick. The small-angle, diffuse x-ray scattering has its greatest width in that direction.

In the fringe micellar structure it is supposed that one and the same molecule can pass through portions that are disordered and through other portions that have crystalline orientation. If the fibers are as in Figure 21·8D, mostly parallel, even though not strictly crystallized, then we have a fiber like that of ramie cellulose. It is at once seen that such a cellulose fiber cannot be stretched appreciably in the *b* or long axis direction because after a very slight elastic elongation the primary
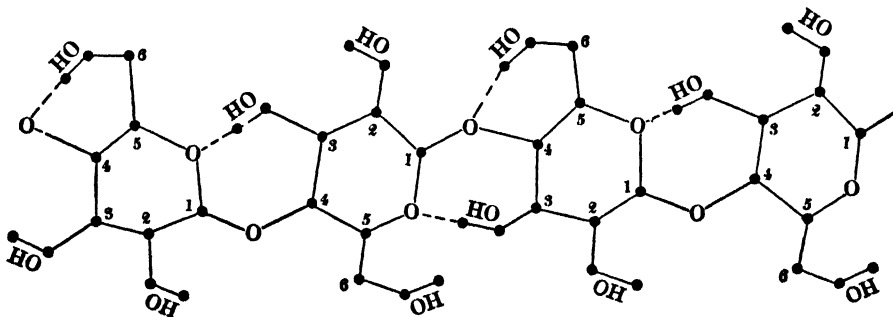
FIGURE 21·7 Arrangement of glucose residues in the fiber axis. Taken from Hermans *et al.*, Text Ref. No. 10.

valences have to be broken. Thus such cellulose exhibits a strong but inelastic structure in the direction of the fibers.

The tensile strength of high polymers of the type of cellulose is surprising until one recalls that it involves the breaking of primary chemical valences. The following Table 21 gives some comparisons which show, for example, that a strand of cotton may be twice as strong as a strand of wrought iron of the same size. Fibers whose high polymer molecules are not oriented along the axis of the fiber will be far weaker. In practice, fibers often slip past each other rather than break their primary valences.

TABLE 21

TENSILE STRENGTHS

Substance	Kilograms per Square Millimeter	Substance	Kilograms per Square Millimeter
Steel wire	64	Rayon	18–20
Steel	50–170	Special rayon, oriented	77
Wrought iron	35	Silk	30–50
Copper	20–50	Nylon, oriented by stretching	50–70
Ramie cellulose	70–95	Rubber	15–20
Flax cellulose	109	Rubber, oriented	60
Cotton (spiral cellulose)	28–77	Thinnest glass fibers	over 1400

It has been stated that resin pressed plywood is the strongest material per unit weight known to engineering science.

Figure 21·2 represents a magnification of 10^8 and explains why cellulose is easily wetted by and swells in water but is not affected by oils or hydrocarbons. It also accounts for the many cellulose derivatives by reaction of the side hydroxyl groups and the possibility of hydrolyzing cellulose to a sugar. Figure 21·5 represents a magnification of 10^7. This explains the double refraction of the crystallized material and its behavior toward dyes. Figures 21·8C and D show a magnification of 10^6 and account for the considerable strength and high elasticity of rayon, its swelling, dyeing, and resistance to creasing. Magnification in a microscope is less than 10^4, but it reveals the fibrillar structure [15] which is responsible for such textile properties as elasticity, softness, and luster. Lesser magnifications are illustrated in Figure 21·1. They show the levelness, soft luster, and homogeneity of the yarn.

Cellulose derivatives are often polymorphic and have longer x-ray spacings than those here discussed. In rayon, for example, there is one of 154 Å and similar unexpected long spacings in other naturally occur-

ring high polymers.[16] Even in cellulose there are several spacings not yet accounted for. Sponsler pointed out that the plant cell wall of cellulose is made up of a hollow ring of perhaps 40,000 layers, the c direction being radial.
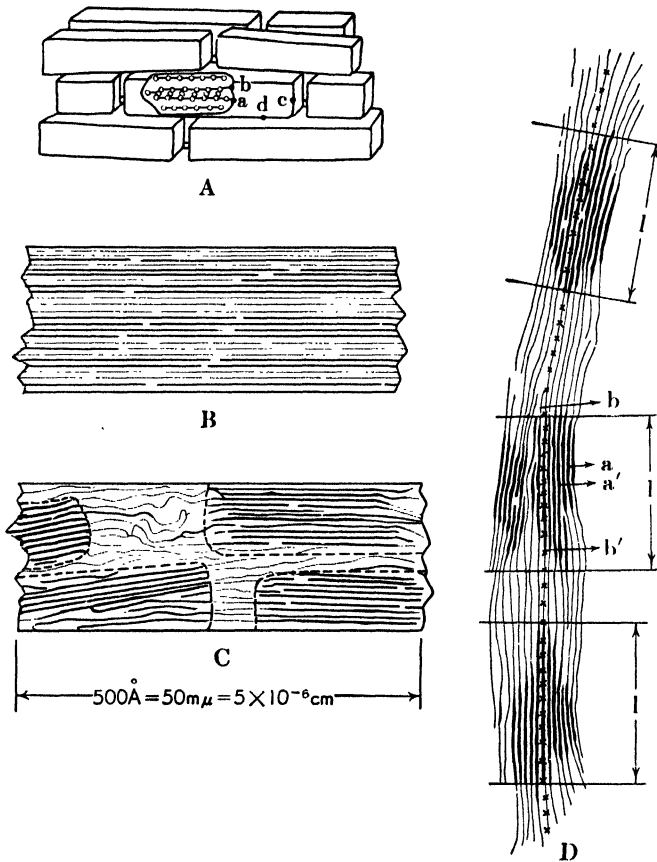
FIGURE 21·8 Schematic representation of several possible crystallite structures of cellulose. The earliest interpretation (A) was that cellulose was built up of separate crystallites. (B) depicts the second, opposite extreme. (C) and (D) depict the presently accepted hypotheses, the network and fringe micellar structures, respectively. Taken from Ott, *Cellulose and Cellulose Derivatives* (see References).

Dr. Wanda K. Farr [17] has isolated uniform cellulose particles about one micron in diameter and has supported the view that the formation of plant cell membranes comes from the incorporation of these particles, which arise in the plastids of living cells. Electron microscopy of pulverized cotton fibers appears to support this view.[18] However, for the cell walls of valonia it is categorically rejected [19] and the continuous

structure indicated in Figure 21·8B is supported rather than the current fringe micelle concept of Figures 21·8C and D.

Cotton fibers differ from those of ramie, although both are nearly pure cellulose in the manner shown in the x-ray diffraction pattern in Figure 21·9. It will be noted that this is a sort of compromise between a fiber and a powder pattern and that diffraction spots have been drawn out into large arcs with an angle of about 57°. This indicates that the cellulose molecules are not parallel to the fiber axis as in ramie but are arranged in a double spiral with an angle of about 57°. This is confirmed by the photomicrograph of a swollen cotton, as in Figure 21·10. Cotton cellulose may, of course, be mercerized to hydrated cellulose with a corresponding change in x-ray diagram and other properties. Banerjee [20] has shown that in some fibers the unit cell is almost parallel to the fiber axis. In others it is arranged spirally, and in still others the arrangement is completely random.

FIGURE 21·9 X-ray diffraction showing natural orientation of mature bundles of cotton fibers. Taken from Sisson, W. A., *Boyce Thompson Contributions*, 9, 242 (1938).

As usual in nature, the surface of a material differs from its bulk. Natural cotton is composed of a primary and a secondary wall. The former forms a thin sheath on the surface of the cellulose fiber and is composed principally of small amounts of cellulose embedded in a membrane of wax and pectic substance.[21] The latter, though present to the extent of approximately only 1%, accounts for 85% of the total acidic groups of fiber and thus greatly influences its properties, as in scouring, dyeing, and finishing. Purified cotton has an isoelectric point at pH 2.5. The cellulose of wood is mixed with lignins, resins, and waxes.

Sorption of water by cellulose [22] and of appropriate organic solvents by cellulose derivatives [23] begins with the characteristic phenomena of true adsorption when exposed to low relative vapor pressure of the solvent, followed by capillary condensation in the pores and crevices opened up by the expansion caused by the sorption. The sorption, of course, includes the new surfaces which the initial sorption creates. Hermans prefers "the point of view that water is dissolved in the amorphous part of the fiber substance and that, in the initial phases of adsorption, formation of true hydrates takes place . . . and may be regarded as an approximate relative measure of the amount of
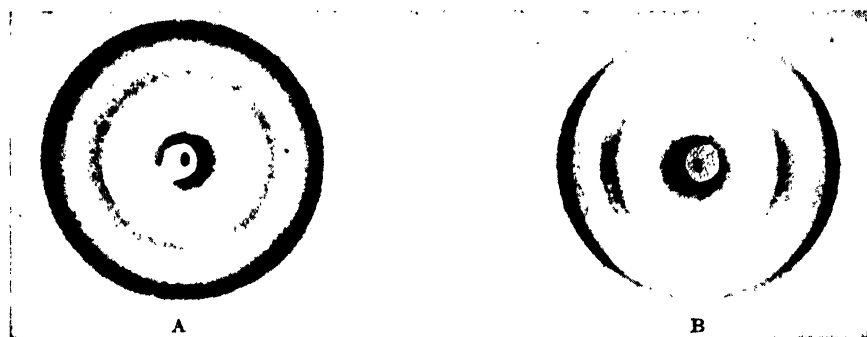
FIGURE 21-10 X-ray diffraction showing natural orientation of young cotton fibers, (A) 15 days, (B) 25 days. From Sisson, W. A., *Boyce Thompson Contributions*, 9, 242 (1938).

amorphous matter in different fibers." [21] The sorptive power of regenerated fibers, which are not hydrate cellulose, is about twice that of native cellulose and exhibits intramicellar swelling.

The uses of cellulose and of cellulose derivatives in fibers, textiles, membranes, and films and in lacquers, paints, and adhesives are too well known to need description in this book. Methyl cellulose may be made water soluble. Recently so-called carboxy methyl cellulose, CMC, in which the hydrogen of the most exposed hydroxyl of each glucose residue is replaced by a $-CH_2COONa$ group, has come into prominence as another water-soluble cellulose derivative, very useful as a thickening agent, and as a protective colloid added to detergents it helps stabilize suspended dirt and avoid its redeposition.

REFERENCES

Cellulose and Its Derivatives

HERMANS, P. H., *Contribution to the Physics of Cellulose Fibres. A Study in Sorption, Density, Refractive Power and Orientation*, Elsevier Publishing Co., Inc., Amsterdam, Brussels, London, New York (1946)

MARK, H., *Chem. Rev.*, 26, 169 (1940)

OTT, E., ed., *Cellulose and Cellulose Derivatives*, Interscience Publishers, Inc., New York (1943)

SISSON, W. A., *Chem. Rev.*, 26, 187 (1940)

Text References

1. DORE, W. H. and SPONSLER, O. L., *Colloid Symp. Monog.*, IV, 174 (1926)
2. HERZOG, R. O., *J. Phys. Chem.*, 30, 457 (1926)
3. HIBBERT, H. and BARSHA, J., *J. Am. Chem. Soc.*, 53, 3907 (1931)
4. Editorial Report, *Nature*, 138, 175 (1936)
5. BRAGG, W. H., *Nature*, 125, 634 (1930)
6. MEYER, K. H. and MARK, H., *Ber.*, 61, 593 (1928); MARK, H. and MEYER, K. H., *Z. physikal. Chem.*, 2B, 115 (1929); MARK, H., *Chem. Rev.*, 26, 169 (1940)

7. ANDRESS, K. R., *Z. physikal. Chem.*, **4B**, 190 (1929)
8. BARRY, A. J., PETERSON, F. C. and KING, A. J., *J. Am. Chem. Soc.*, **58**, 333 (1936)
9. MCBAIN, J. W. and STUEWER, R. F., *J. Phys. Chem.*, **40**, 1157 (1936)
10. HERMANS, P. H., DE BOOYS, J. and MAAN, C. J., *Kolloid-Z.*, **102**, 169 (1943)
11. FANKUCHEN, I. and MARK, H., *Record of Chemical Progress*, **4**, 57 (1943)
12. HERMANS, P. H. and WEIDINGER, A., *J. Appl. Phys.*, **19**, 491, 505 (1948)
13. MYSELS, K. J. and MCBAIN, J. W., *J. Phys. Coll. Chem.*, **52**, 1471 (1948)
14. KRATKY, O., *Kolloid-Z.*, **96**, 301 (1941)
15. KINSINGER, W. C. and HOCK, C. W., *Ind. Eng. Chem.*, **40**, 1471 (1948)
16. CLARK, G. L. and PARKER, E. A., *Science*, **85**, 203 (1937)
17. FARR, W. K., *J. Phys. Chem.*, **42**, 1113 (1938); SISSON, W. A., *Chem. Rev.*, **26**, 187 (1940)
18. BERKLEY, EARL E., A.C.S. meeting, April, 1942
19. PRESTON, R. D., REED, E. and MILLARD, A., *Nature*, **162**, 666 (1948)
20. BANERJEE, B. K., *Indian J. Phys.*, **30**, 259 (1947)
21. SOOKNE, A. M. and HARRIS, M., *Am. Dyestuff Reptr.*, **30**, 29 (1941)
22. SHEPPARD, S. E. and NEWSOME, P. T., *Ind. Eng. Chem.*, **26**, 285 (1934)
23. MCBAIN, J. W., GOOD, S. J., BAKR, A. M., and DAVIES, D. P., *Trans. Farad. Soc.*, **29**, 1086 (1933)
24. HERMANS, P. H., *Contribution to the Physics of Cellulose Fibres*, op. cit., p. 54

OTHER NATURAL POLYSACCHARIDES

Starch

Starch [1] is one of the best known examples of naturally occurring polysaccharides. Its chemical constitution is based on glucose residues, as in the diagram in Figure 22·1. The distinction is that in cellobiose there are beta-glucosidic linkages, and therefore the cellulose molecule would tend to be linear, whereas in amylose and starch the glucose residues have an alpha-glucosidic linkage which leads to a spiral structure.

There are two main different components in starch, amylose and amylopectin, with very different molecular structure, although both are made up of maltose. Both have been synthesized by means of enzymes.[2] A portion of the amylose may be extracted by means of hot water, in which amylopectin is insoluble. The amylose consists of one long, unbranched series of 50 or more maltose residues connected end to end. According to the ultracentrifuge it has a molecular weight of 88,000. Its length has been determined by the end-group analytical method. As would be expected, the amylose derivatives form tough fibers and films. Pure amylose has been successfully crystallized, and the crystals are found to be made up of helically wound molecules with six glucose residues per spiral. Blue starch iodide consists of such spirals, which are filled with iodine atoms laid end to end for which there is just room within the helix.
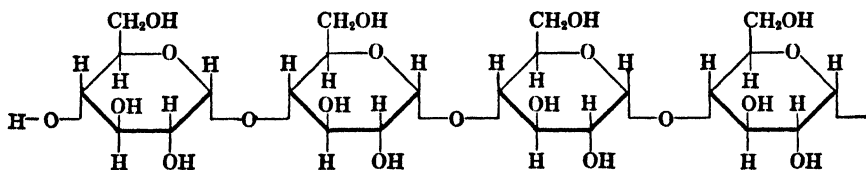
FIGURE 22·1 Diagram showing the structural arrangement of the starch molecule.

Starch itself gives chiefly a diffuse halo with x rays and if hydrated also shows a few indications of some crystalline or micellar structure. The hydrate contains about 10% water, or one mole water to each glucose residue. The amylopectin is amorphous, and its derivatives give brittle films and never fibers. The end group analytical method gives one end group for every 10–15 maltose residues. However, this is not due to small molecular weight. On the contrary, the molecular weight must be at least 4500 [3] but according to ultracentrifuge evidence it lies between 400,000 and 700,000.

The explanation is that the amylopectin is highly branched, in contrast to amylose. Meyer [4] gives reasons for the interpretation as highly irregular branching, whereas Haworth [5] suggested a laminated branching structure in which there is a 1–6 glucosidic linkage between the glucose at each branch and the middle of the next branch. There appear to be no cross-linkages. Ultracentrifuge evidence [6] indicates that cornstarch consists of one quarter amylose, one quarter amylopectin just described, and, to the extent of one half, another amylopectin which is quite heterogeneous but whose molecular weight ranges from 3 million to 11 million according to the ultracentrifuge. Such giant branching molecules cannot wholly crystallize, and to them and their numerous spiraling branches must be attributed most of the valuable properties of starch. Partial crystallization is said to be involved in the staling of bread.

Only the longer branches of amylopectin are long enough to produce even reddish colors with iodine. Degradation by heating, acids, or enzymes progressively breaks down amylopectins to produce smaller branched fragments only the largest of which still give a red color with iodine. The ultimate product is glucose. The enzyme amylase, according to ultracentrifuge measurements, has a molecular weight of 45,000. [7] Seeds, tubers, and pith contain starch granules of the order of several microns, arranged radially in concentric layers to make up the grain of starch.

Starch grains are insoluble in water at room temperature and the suspension is dilatant, but they form a transparent, highly viscous and thixotropic paste or jelly with boiling water which can be used as an adhesive but is preferably first partly degraded. Boiled starch paste owes its thickness to the amylopectin because amylose does not form such a paste. The whole structure swells and becomes transparent. The temperature at which the starch "paste" or jelly forms is greatly affected by the presence of electrolytes, and many regularities, such

as lyotropic series and Traube's rule with regard to effect of molecular weight within a homologous series, have been observed although other effects also appear.[8] Different sources of starch give quite different products with different proportions of amylose and amylopectin. Boiled starch paste is of value in clogging and sealing off with a thin film of highly viscous amylopectin the walls of oil wells in drilling. Similarly, in laundering, only a thin superficial coating is involved.

Similar effects of additives are observed in connection with swelling and viscosity. Extremely dilute soap lowers the viscosity; less dilute raises it.

There has always been a tendency in the field of colloids to mention only the main constituent, as with colloidal "gold" or "platinum," although the chief properties are determined by the stabilizing agents or the surfaces. Possibly this is here also the case. Dextrin and starch, unless especially treated, contain phosphoric acid to the extent of one mole per 10,000 grams. In addition they always [9] contain organic acid groups, supposed by Pauli to be uronic acid, but also fatty acids, probably as lecithin or other phosphatides. Samec designated such starches colloidal electrolytes. Dextrin and starch readily form derivatives analogous to those of cellulose with alkalis, with nitric and fatty acids, and by methylation. The osmotic pressures [10] of methylated potato starches and dextrins correspond to particle weights of as low as 2910 for a dextrin, 38,200 for a maize (corn) starch, and 124,000 for a potato starch.

Glycogen

Glycogens are the readily available sugar reserve of animals, as starch is that of plants. They occur in yeast, liver, muscle, and mussels. They are closely related to starch, give a brown or violet color with iodine, but do not form granules. Like starch they are based upon maltose residues, and like amylopectin they are highly branched but the branches contain less than half as many maltose units. The osmotic pressure [11] of acetylated glycogens corresponds to a particle weight of several million, that of methylated glycogen to less than a million. These values are far larger than would be expected from viscosity data.

Glycogen and starch can be synthesized by the enzyme phosphorylase. Enzymes that attack starch also reduce glycogen to dextrins and maltose. Starch and glycogens form similar derivatives. In nature, glycogens seem often to be combined with protein.

Other Polysaccharides Related to Cellulose

Xylan and *polyglucuronic acid* differ from cellulose, with which they commonly occur, in having only a hydrogen atom or a carboxyl group, respectively, in place of the CH_2OH side group of the glucose residues. Needless to say the carboxyl group leads to many derivatives and governs much of the behavior. Xylan is extracted from grass. It is of relatively low molecular weight for a polymer, containing perhaps only seventy to eighty sugar residues. It is soluble in dilute alkali and gives solutions of relatively low viscosity. Xylan consists of branched chains which, however, terminate in arabofuranose groups.

Chitin forms the hard shell of crustaceans, insects, and fungi and has the structure of cellulose except that an acetylated amine group replaces the hydroxy group on the second carbon atom of the glucose residues. Its x-ray repeating spacing is therefore 10.4 Å. It is soluble in most cellulose solvents. The molecular weight is of the same order as that of cellulose.

Lichenins occur in lichens and Iceland moss. They consist of straight chains of cellobiose residues of particle weight from eleven thousand to one hundred and eighteen thousand. They dissolve in hot water and form a jelly on cooling. They must differ in some way from the detailed structure of cellulose.

Inulin is the sugar reserve of the tubers of the compositae, such as dahlias and Jerusalem artichokes. It has been thought to be dissolved in the cell sap. At 20° C its solubility in water is only 0.01%, but on warming it becomes freely soluble and can crystallize. Its molecular weight is of the order of only six thousand. The structural unit, unlike that of cellulose, is unique in being fructofuranose. End group analysis indicates a straight chain of less than thirty fructose residues. Osmotic pressure measurements [12] of methylated and acetylated inulin give 6210 and 8880 as the particle weight. The viscosity value is very low.

Besides those mentioned above, many less well studied polysaccharides often occur, such as the *mannans;* for example, that from the tubers of Konjaku is water soluble and may be made into fibers with a repeating spacing of 8.5 Å.

Lignin is not a polysaccharide, but it commonly occurs closely intermixed with cellulose in wood, straw, jute, etc. Its molecular weight is unknown but may be in the neighborhood of 7000, which may artificially be halved or increased fourfold. [13] "The axial ratio of the 7000 molecule is about 5." Possibly it constitutes a three-dimensional network. Its unit is thought to be a substituted phenyl propane and con-

tains methoxy groups An unoriented x-ray spacing of 4.2 Å belongs to lignin together with equatorial reflexions of 9.7 Å and 14.6 Å. Extracted soluble fractions of lignin give a molecular weight of 3900 by osmotic pressure and rise of boiling point; that is, four times the empirical weight of approximately 900; but 10,000 by diffusion and dielectric constant, the latter ascribed to influence of shape.

Pectins [14]

Not glucose or maltose, but galactose, arabinose, and galacturonic acid and some rhamnose constitute the structural units of the pectins. These are characterized by their power of swelling greatly in water to form thick mucilaginous masses and jellies. They are of widespread occurrence in substantial proportions in cell walls and fruits and roots of plants. They are extracted by boiling water. The free acids are insoluble in acid solutions and do not gel; neither do the methyl esters. Insoluble pectin is at least partly the calcium or magnesium salt of pectinic acid.

Pectins are mixtures consisting of linear chains with side groups of arabans and galactans. The main constituent is pectic acid, which is a polygalacturonic acid of the formula: [15]
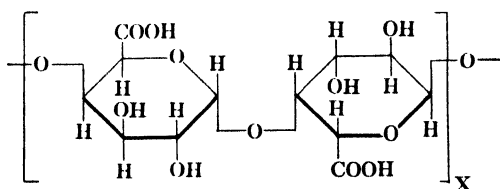
It is often partly methylated, and its molecular weight may be one hundred thousand or more. Its equivalent weight is about 205, which is somewhat more than would be calculated from the repeating part of the above formula.

Sodium pectate forms fibers [16] with an identity period of 13.1 Å. The distance between chains increases from approximately 6.6 Å to 6.94 Å on methylation, or from 6.25 to 7.1 Å with uptake of water. The crystalline portion is hydrated. Stuewer [17] found by ultrafiltration and by swelling, using reference substances, that the hydrate water amounts to 0.25, 0.21, 0.38, and 0.35 gm of water to 1 gm pectin, pectic acid, sodium pectate, and calcium pectate, respectively. The amount of water taken up by the crystallites between chains is 24%. [18]

The osmotic pressure of this colloidal electrolyte, taken as if it were a nonelectrolyte, would correspond to 13 units in the chain. The

conductivities, degrees of dissociation, and osmotic coefficients of apple pectin, pectinic acid, and of various pectinates have been measured. The free acid has a dissociation constant of the order of 10^{-4}, and the pectinates behave as expected for colloidal electrolytes. The solutions show streaming double refraction. "The aggregation tendencies of pectin depend principally upon lateral forces acting between the pectin molecules. These forces arise for the most part from the carboxyl groups along the polygalacturonide chain," and "depend greatly upon the ionic equilibria."

Commercial pectins are used in the production of jellies, and their preparation involves some degradation of the molecule. With orange peel pectin an optimum formula contains only 0.40 gm pectin with 62.58 gm sugar per 100 cc at pH 2.60, the jellying being reversible with heat. A more general optimum for pectins lies at pH 3.5. However, a more extensive study [19] shows that these optimum recipes have no general significance and refer only to specific instances of jelly making. The jelly is a brush-heap structure, and there is a balance between the solubilizing groups and the tendency of the insoluble portions of the molecule, such as the methylated carboxyl groups, to adhere by being thrust out of solution. An optimum degree of methylation lies at about 12% methoxyl. Alcohol and sugar reduce the solubility, and therefore the latter promotes jellying.

Plant Gums [20]

Gums and mucilages are likewise highly swelling, high polymers with enough dissociable groups to make them colloidal electrolytes. They contain residues of d-glycuronic acid, d-galactose, l-arabinose, and in lesser amounts, d-mannose, d-xylose, and l-rhamnose and occur in nature largely in the form of their calcium and magnesium salts. They do not form fibers. They are of variable composition as regards the proportion of the constituent residues mentioned. Some of the more important gums are locust bean, tragacanth, quince seed, linseed, karaya, and acacia. The latter is an especially effective emulsifying agent. Tragacanth is a good thickening agent.

Gum arabic, obtained from acacias, is the best studied example. Its equivalent weight is nearly 1200, somewhat greater than that of gum tragacanth, 1050, but measurements of osmotic pressure [21] show that its molecular weight is several hundred thousand, which would correspond to about 240 acidic groups. The free acid has a dissociation constant of the order of 10^{-3}, so that in 1% solution it would have 200

undissociated and 40 dissociated groups. The viscosity of the sodium salt is low at pH 2 and pH 12 but passes through a high flat maximum between pH 5 and pH 9. Added salts lower the viscosity, polyvalent cations having the greatest effect. The protective action of the free acid is much less than that of the salts.

Here may be mentioned one of the unsolved problems of colloidal electrolytes: the discrepancy between ion concentrations as determined by electromotive force and by other thermodynamic methods and to a lesser extent the quantitative divergence between such data and those of conductivity and the so-called Hammarsten effects.[22]

Marine Polysaccharides

Agar (or *agar-agar*) and *carrageen* mucilage are well-known colloidal electrolytes, remarkable for their thickening properties, which agar possesses in far higher degree than gelatin. Agar is derived from *d*-galactose ending in *l*-galactose esterified on its sixth carbon with a sulfonic acid group, giving an equivalent weight of about 9000. The effects of electrolytes upon its swelling have been much studied. These are similar to those just described for gum arabic. In concentrated solutions the lyotropic series is evident.

Carrageen mucilage [23] is the calcium salt of a carbohydrate (galactose, etc.) ethereal sulfate containing very much more sulfate than does agar; the ash contains 24% sulfate. Some of the viscosity is attributable to the divalent calcium bonding adjoining chains.

Alginic acid constitutes a large proportion of the dry weight of certain seaweeds. It is noteworthy among polysaccharides in that it appears to be built up entirely of *d*-mannuronic acid residues [24] according to the formula:
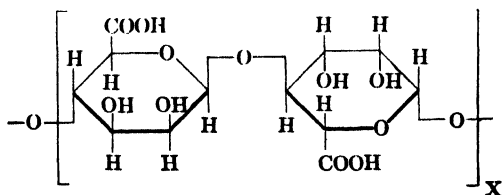
This formula should be compared with those of cellulose and pectic acid. The unit cell contains 4 mannuronic acid residues and 4 water molecules, but the fibers have a repeating pattern of only 8.7 Å, showing a wavy structure of the sugar residues as compared with cellobiose. The other two dimensions are 7.75–8.7 Å and 10.6. Å.[25] Like so many

other organic acids such as the fatty acids and the bile acids, the free acid and the calcium salts are insoluble in water, whereas the sodium salts are soluble as colloidal electrolytes. The viscosities of the alginates indicate molecular weights of the order of 15,000. The osmotic pressure gives values from 48,000 to 185,000. The viscosity is at a maximum near pH 7.

Commercial preparations of alginates have assumed great importance in recent years on account of their many applications as thickening agents. They are widely used in ice cream and other dairy products.

The most recent development is that of "seaweed rayon." [26] This is spun as sodium alginate and converted to alginic acid or calcium alginate, then oiled with a detergent to prevent adhesion and to give good "handle." Alginic acid yarn is soluble in alkali or soap solution but may be cross-linked with formaldehyde and then, like chromium and beryllium alginate rayons, it becomes resistant. The usual yarns are calcium alginate, but the solubility and disappearing properties of alginic acid and sodium alginate have been turned to account in producing special textiles.

REFERENCES

Other Natural Polysaccharides

HAWORTH, W. N., "A Presidential Address Delivered before the London Chemical Society," *J. Chem. Soc.*, 543 (1946)

KARRER, P., *Einführung in die Chemie der Polymeren Kohlenhydrate*, Akademische Verlagsgesellschaft, Leipzig (1925)

MEYER, K. H., *Natural and Synthetic High Polymers*, Interscience Publishers, Inc., New York (1942)

VALKÓ, E., *Kolloidchemische Grundlagen der Textilveredlung*, Julius Springer, Berlin (1937)

Text References

1. SAMEC, M., *Kolloidchemie der Starke*, Steinkopf, Dresden and Leipzig (1927), pp. 646–682
2. HAWORTH, *loc. cit.*
3. HIRST, E. L. and YOUNG, G. T., *J. Chem. Soc.*, 1471 (1939)
4. MEYER, K. H., *Naturwissen.*, 28, 564 (1940); *Nature*, 160, 900 (1947)
5. HAWORTH, *loc. cit.;* also *Nature*, 160, 901 (1947)
6. FOX, T. G., BECKMANN, C. L., A.C.S. meeting, 1943
7. DANIELSSON, C. E., *Nature*, 160, 899 (1947)
8. KATZ, J. R., MUSCHTER, F. J. F. and WEIDINGER, A., *Biochem. Z.*, 257, 397, 404 (1933); 261, 19, 48 (1933); 271, 54 (1934)
9. PAULI, W. and SZPER, J. and ST., *Kolloid-Z.*, 82, 335 (1938); SZPER, J. and ST., *Trans. Farad. Soc.*, 35, 667 (1939)
10. CARTER, S. R. and RECORD, B. R., *J. Chem. Soc.*, 660–675 (1939)
11. CARTER, S. R. and RECORD, B. R., *loc. cit.*

12. CARTER and RECORD, *op. cit.*, p. 665
13. GRALÉN, N., *J. Coll. Sci.*, **1**, 453 (1946)
14. SPEISER, R., *J. Polymer Sci.*, **2**, 281 (1947)
15. HIRST, E. L., *J. Chem. Soc.*, 70 (1942)
16. PALMER, K. J., MERRILL, R. C., OWENS, H. S. and BALLANTYNE, M., *J. Phys. Coll. Chem.*, **51**, 710 (1947)
17. STUEWER, R. F., *J. Phys. Chem.*, **42**, 305 (1938)
18. PALMER, K. J., SHAW, T. M. and BALLANTYNE, M., *J. Polymer Sci.*, **2**, 318 (1947)
19. OLSEN, A. G., STUEWER, R. F., FEHLBERG, E. R. and BEACH, N. M., *Ind. Eng. Chem.*, **31**, 1015 (1939)
20. MANTELL, C. L., *The Water-Soluble Gums*, Reinhold Publishing Corp., New York (1947)
21. OAKLEY, H. B., *Trans. Farad. Soc.*, **31**, 136 (1935); **32**, 1360 (1936)
22. PAULI, WO. and RIPPER, E., *Kolloid-Z.*, **62**, 162 (1933); MCBAIN, J. W., *J. Am. Chem. Soc.*, **57**, 1916 (1935)
23. PERCIVAL, E. G. V. and BUCHANAN, J., *Nature*, **145**, 1020 (1940)
24. KRINGSTAD, H. and LUNDE, G., *Kolloid-Z.*, **83**, 203 (1938); HIRST, E. L., JONES, J. K. N. and JONES, W. O., *Nature*, **143**, 857 (1939); *J. Chem. Soc.*, 1880 (1939)
25. ASTBURY, W. T., *Nature*, **155**, 667 (1945)
26. SPEAKMAN, J. B., *Nature*, **155**, 655 (1945); CADY, W. H., *Amer. Dyestuff Reptr.*, **37**, 283 (1948)

NATURAL AND SYNTHETIC RUBBER

Rubber Latex

Rubber occurs in nature in the form of latex in the sap of *Hevea brazilien-sis* and a number of other trees and plants such as *Castilloa elastica*. A hevea latex is a white cream of latex particles suspended in water. It contains 30–40% rubber, and its properties are due largely to the stabilizing agent, which is mainly protein and some fatty acid amounting to 5% of the weight of the rubber. The amino acids of the protein have been quantitatively determined.[1] There are also present some resins and phytosterol. The rubber particles have an interesting typical shape illustrated in Figure 23·1. These particles vary in diameter from 5000 to 30,000 Å and there are about 50 million of them per cubic millimeter. The larger particles can be dissected by means of a micromanipulator, and it can be shown that the interior consists of liquid or sol rubber surrounded by a thickened, tough gel of rubber coated on the outside with the stabilizing layer of protein. Both have the same composition, but the sol rubber is soluble in ether, whereas the gel rubber is not, but merely swells. Gel rubber dissolves almost completely in benzene containing butyl alcohol. Gel rubber may contain oxygen to the order of 0.01%, which would be sufficient to explain cross-linking and resistance to solution. Here again we have an example of the importance in colloid science of minor surface groupings in determining some of the major properties of the material.

The properties of latex or diluted latex can be predicted from the properties of protein. For example, the particles will be independent, and therefore dilute latexes will have a viscosity only somewhat greater than that predicted by the Einstein formula. This is in great contrast to solutions of rubber, to be discussed later, where the viscosity of a

5% solution of rubber in benzene is 100 times greater than that of a 5% rubber latex. *Hevea* latex has a pH of 6.0, or, in the absence of carbon dioxide, 6.25, and it is negatively charged. There will be an isoelectric point which, for *Hevea*, occurs at pH 4.8, but for different species of *Euphorbia* is nearer pH 4.[2] Serum from *Hevea* latex has an isoelectric point of 4.7.

Away from the isoelectric point, electrophoresis is observed, and this has been developed into a method of manufacturing high grade rubber articles by electrodeposition. Coagulation will occur at the isoelectric point upon the addition of acids, such as hydrochloric or acetic, or smoke. If enough hydrochloric acid is added to exceed pH 3 there is no coagulation and the particles are charged positively, but a much larger excess of hydrochloric will coagulate them again.

Recently it has been found that in addition to the rubber particles, there are also viscous, colloidal bodies soluble in ammonia and alkalies which contribute noticeably to the properties of the latex.[3]
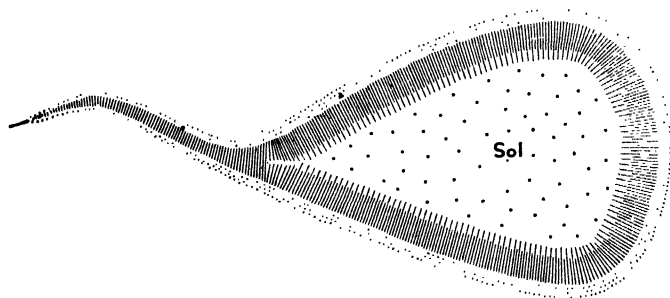
FIGURE 23-1 Rubber latex particle.

If a latex particle on the point of a needle be placed in benzene, it swells to a globule of several times its diameter and then bursts, the contents dissolving in the benzene and the skin remaining on the needle. If the particle is vulcanized, it first becomes nearly globular and then solidifies. Latex is stabilized by the addition of ammonia. Its creaming can be accelerated by the additions of such colloids as gum tragacanth or Iceland moss. Latex particles may be protected with soaps and concentrated up to 70% rubber. On the other hand, soaps in intermediate range, depending upon the magnesium and calcium contents of the latex, can coagulate the latex, and the formation of the calcium-magnesium soaps in *Hevea* latex can account for spontaneous coagulation.[4]

Articles may be made from latex by dipping to make gloves, balloons, etc., using coagulants or thixotropic dips. Microporous rubber

is made by foaming latex; rubber threads are made by extrusion into a coagulating bath. Latex may be cast in molds as a thickened gel and then vulcanized. Its adhesive and plastic properties are of value in the leather industry, and for sealing tins in canning.

The coagulation of latex by electrolytes follows in general the universal rules already discussed in Chapters 12 and 18, polyvalent cations being most effective. The stability of the latex is diminished by bacterial decomposition of the protein. Indeed, fresh *Hevea* latex can be kept fluid for only a few hours without preservatives. The object of coagulation is not merely to separate out the rubber but also to give it a suitable structure. Often the freshly coagulated latex has not coalesced and the process may then be reversed, but in time the particles adhere and coalesce. Different processes retain more or less of the other constituents of the latex and therefore differ greatly.

Milling is important in the treatment of coagulated rubber in order to break up the particles and bring about their complete coalescence and also to reduce the viscosity of the resulting rubber. The milling is carried out by polished revolving rollers rotating at different speeds so that the rubber is torn and "masticated." If this is done in an atmosphere of nitrogen, the expected softening of the rubber does not occur. This shows that the process is not merely one of rupturing the rubber globule and releasing the softer, semi-internal rubber, or a depolymerization by reduction of the rubber molecules to shorter lengths, but that the breakdown of the rubber is accompanied by oxidation induced by static electricity. Hence peroxides, mercaptans, and oxidized mineral oils are added, which permit attaining properly plasticized rubber with a material reduction in time, power, and cost.

It is perhaps characteristic of colloid technology that we speak of rubber but do not mention the important although little known role of the stearic acid, zinc oxide, and zinc stearate that is employed during mastication. Masticated *Hevea* rubber consists of about 94.5% caoutchouc and 5.25% resins, proteins, and fatty acids, with 0.25% ash. The proteins exert noticeable effects in adsorption, even in finished rubber.

Domestic Sources of Rubber

Minor amounts of natural rubber are obtainable from plants such as guayule, *Cryptostegia*, kok-saghyz, goldenrod, milkweed, pinguay, and rabbit brush.[5] Of these, guayule has been most exploited in the United States and kok-saghyz in the U.S.S.R.

Guayule plants resemble sagebrush in appearance and grow wild in Mexico but have been cultivated in California. The rubber is enclosed in the individual cells of the fibrous shrubs, which therefore have to be comminuted in pebble mills to release the rubber, which floats on water, allowing the water-logged plant material to sink. The production consists on the average of 70% rubber hydrocarbon, 20% resin, and 10% of material insoluble in a mixture of benzene and acetone, chiefly cellulose and lignin. The resin may be largely removed by a retting process. The rubber can be recovered by alkali-alcohol extraction or as latex.[6] Guayule rubber is particularly useful as the inner swelling layer of airplane fuel tank cells. It adds tack and other desirable properties when blended with other natural and synthetic rubbers. Unless provided with an antioxidant, it slowly turns to a viscous black tar after about a year, and this tar forms an excellent protective colloid. The amount of rubber in the guayule shrub has been increased by plant breeding to as much as 25% of the dry weight of the fiber. The details of the completely mechanized production of guayule are fascinating.

X-Ray Behavior of Rubber [7]

As was shown by Faraday in 1826, natural rubber is essentially polyisoprene with the empirical formula $(C_5H_8)_x$. Isoprene,

$$
\begin{array}{ccc}
CH_2 & & CH_2 \\
\diagdown & & \diagup \\
& C\!-\!CH & \\
& | & \\
& CH_3 & \\
\end{array}
$$

forms the repeating unit,

$$
\begin{array}{ccc}
CH_2 & & CH_2 \\
\diagup & & \diagup \\
\diagup\; C\!=\!CH \;\diagdown & \\
& | & \\
& CH_3 & \\
\end{array}
$$
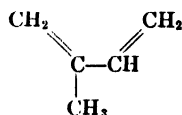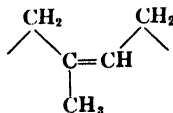
The corresponding *trans* formula is that of gutta percha, balata, and chicle.

It was an outstanding event for the study and understanding of high polymers when Katz in 1925 [8] showed that rubber which has been stretched by more than 75% produces a fiber diagram when examined by x rays. Previous workers had obtained only the typical, amorphous diagrams of unstretched rubber as illustrated in Figure 23·2, as compared with the fiber diagram of stretched rubber in Figure
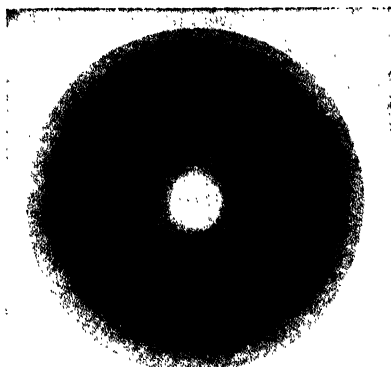
FIGURE 23·2 X-ray diagram of unstretched, nonfrozen rubber (smoked rubber). Taken from Meyer, *Natural and Synthetic High Polymers* (see References).
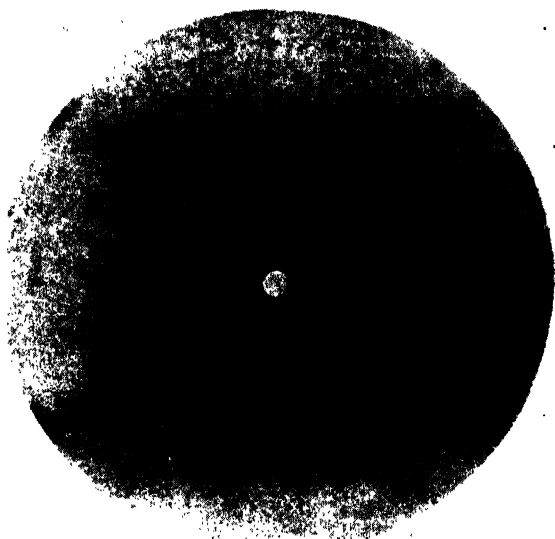
FIGURE 23·3 Fiber diagram of stretched rubber. Preparation inclined at 8° to the normal. Taken from Meyer, *Natural and Synthetic High Polymers* (see References).

23·3. As the stretching is increased the amorphous halo gets weaker and weaker and the fiber pattern becomes more and more definite. The spacings are not changed by further stretching, only the intensity. The stretched rubber is therefore at least partially crystallized, whereas unstretched rubber has a liquid structure. The size of the crystalline regions according to the width of the diffraction lines is of the order of 600 Å in the direction of stretching and 500 Å and 150 Å in the width and thickness. Electron diffraction patterns of unstretched and stretched gutta percha are given in Figures 23·4 and 23·5.

On being kept for a long time at low temperature, stretched rubber, "frozen" rubber, becomes turbid and progressively crystallized. On release it remains in this form at low temperature, "racked" rubber, but on warming snaps back to approximately the original dimensions. If the rubber is merely chilled without stretching, it gradually crystal-

lizes and with x rays produces a powder diagram, as in Figure 23·6. The crystalline state is that which is most stable below about 20° C. Crystallization becomes more complete and stable for many years and may account for the hardening of rubber articles. Whitby found a transition temperature of 43·5° C for a sample of rubber frozen for thirty years, but the x-ray spacings are the same as those of freshly frozen rubber. In spite of long time experiments, true equilibrium is apparently never attained. Both with stretched and with frozen rubber only a minor fraction ever crystallizes. At very low temperatures a "second order" transition occurs, and both crystalline and amorphous states give place to a hard, glassy state.

The repeating unit of rubber is 8.2 Å. That of β-gutta percha is 9.54 Å. The gutta percha, being a *trans* form of polyisoprene, forms comparatively straight chains which lend themselves much more readily to crystallization so that the crystalline form is stable up to 65° C. The phase diagram for the two crystalline forms and the amorphous or liquid form of gutta percha has been given by Hauser and Susich.[9] Various unit cells have been suggested for crystallized rubber but complete agreement has not yet been attained. It is at once clear
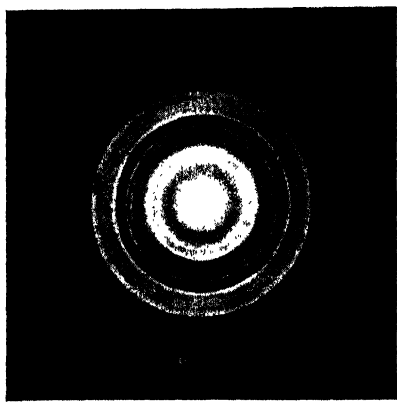
FIGURE 23·4 Powder diagram of unstretched film of gutta percha. Taken from Germer, L. H., *J. Appl. Phys.*, 9, 145 (1938).
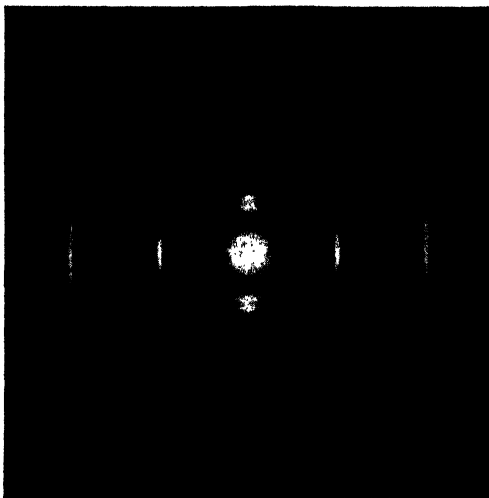
FIGURE 23·5 Fiber diagram of stretched gutta percha. Taken from Germer, L. H., *J. Appl. Phys.*, 9, 145 (1938).

that such crystallization of straight chains, just like that of cellulose, does not give appreciable elastic stretching in the direction of the fiber axis; whereas the way in which the rubber molecules can spontaneously depart from ordered structure by contracting into other shapes is in accordance with the well-known elasticity of rubber. In the amorphous form the various molecules are free to move and take up different configurations and to slip past each other.
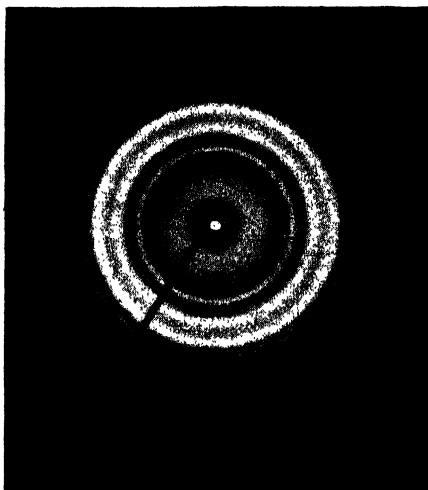
FIGURE 23·6 Powder diagram of frozen rubber. Taken from Gehman, Text Ref. No. 7.

Current concepts of the structure of partly crystallized, long-chain polymers before and after stretching and of the molecules of an elastic polymer such as rubber are given in Figure 23·7.

The fundamental reason for the elasticity of rubber is to be found in the thermal motion of the chain atoms of the very long polymer. They will take up a certain normal distribution in much the same way as the molecules of a gas acquire a definite distribution of velocities. Any deformation of rubber disturbs this normal distribution and hence leads to a less probable and less stable state. As in gases, the tension required for extension is proportional to the absolute temperature and its evolution of heat proportional to the work done. By thermodynamic reasoning [10] the probability has been related to the entropy and the work of deformation and from this the stress-strain relationships can be derived. An alternative approach has been given by Kubo.[11]
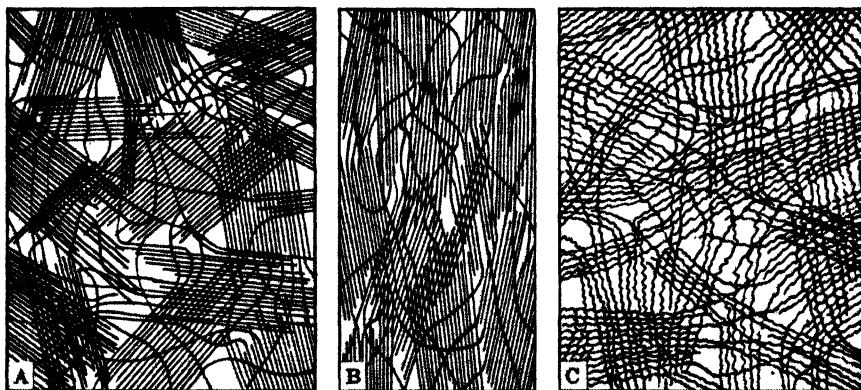
F I G U R E 23-7 Current concept of crystallized elastic polymer. (A) Crystalline, long chain polymer. Lines represent molecules, groups of parallel straight lines represent crystals. (B) The same, cold-drawn. Crystals apparently parallel. (C) Amorphous rubberlike polymer. Wavy lines represent wriggling molecules. Taken from Bunn, *Proc. Roy. Soc.*, **180A**, 95 (1942).

It should be remembered here that stretched rubber occupies the same volume as the original, unstretched rubber. The properties of rubber depend upon both strain and the rate of strain.

Roberts [12] maintains that the elastic properties of rubber are dependent upon the presence of a substance of the formula $C_{80}H_{130}(OH)_2$.

Vulcanization

Vulcanization makes rubber more elastic whereas unvulcanized rubber is permanently deformed. Vulcanization is not completely understood, but an important function is to create sulfur bridges between adjacent molecule chains, thus fixing their position relative to each other and producing a three-dimensional framework. When the links are few, the molecules are prevented from slipping past each other but still retain their power of retracting into coiled and twisting forms, thus making the rubber elastic.

Trapeznikov [13] in studying the exceptionally high elastic constants of very thin films obtained evidence for the existence of micelles in rubber sols and points out that in elastic deformation the micelles must not slip past each other. Gehman and Field [14] from comparison of light scattering and viscosity behavior conclude that: ". . . in rubber solutions, the rubber molecules exist in clusters or groups, the size, shape and 'interlocking' of which depend upon the molecular forces in the solution." Unvulcanized rubber is softened by heat, is "thermoplastic," and is then very easily handled and molded. On the other hand,

vulcanized rubber is no longer thermoplastic and if heated, it may scorch or burn without softening. It is of great practical importance that the rubber can be processed into any desired shape and then made permanent by vulcanization.

The formation of cross-linkages is a vital question in the discussion of high polymers. Such cross-linkages are almost wholly absent from the high polymers discussed in the previous chapters but are of the greatest importance in discussing proteins and especially resins and plastics. The cross-linking is progressive as more and more sulfur is used for bridging, until finally a hard, three-dimensional, brittle network of ebonite is produced. Vulcanization also affects the appearance of the x-ray fiber diagram, which requires very great stretching for its production, say over 250%. Sol rubber produces no x-ray evidence of crystallinity, even when stretched by 1000%, until it has been vulcanized.

Fillers and Reinforcing Agents

Suspending suitable solids, of which carbon black is the outstanding example, in rubber greatly improves the mechanical properties of the rubber, as in tires. Thus the very great tensile and abrasive strength of the rubber used for tires is produced by the addition of large amounts of a very active kind of carbon black. The toughness and resistance to wear and to cracking depend very much upon the fineness and the suitability of the particular carbon black. The carbon black has specific affinity for the rubber and is firmly attached, serving to anchor and also orient the surrounding molecules of rubber in which it is embedded. Similarly zinc oxide is a true reinforcing agent, but many other solids are added to commercial rubber simply as diluents or fillers. Some coloring agents can act as reinforcing agents. Rehbinder, Ab, and Weiler [15] showed that active fillers produce well-defined, continuous structure throughout rubber, whereas inactive fillers do not.

Stamberger [16] compared the affinities of various solvents by dissolving strongly masticated, unvulcanized rubber in gasoline and after covering this with a layer of pure gasoline, noting how much the powder was carried along with the rubber into the gasoline. For example, the amount of gas black or zinc oxide so carried was about fifty times greater than that of clay. Many other ingredients are added to rubber, such as antioxidants, accelerators for vulcanization, and softeners that facilitate the slipping of unattached parts of molecules past each other. A high degree of dispersion of a filler gives greater tensile strength.

Molecular Weight of Rubber

The molecular weight of rubber according to osmotic pressure ranges from 80,000 to 300,000, depending upon the degree of mastication. Diffusion indicates much higher values, from 830,000 to 970,000, but excessively dilute solutions, much less than 0.1%, must be used to achieve independence of the molecules.[17] The ultracentrifuge has given a molecular weight of 430,000–435,000. The relation between intrinsic viscosity and molecular weight obtained from lowering of freezing point is $[\eta] = 5.02 \times 10^{-4} \times M^{\frac{2}{3}}$. Milled and unmilled rubber in toluene does not pass through ultrafilters with 500 Å pores but does through filters of 50,000–300,000 Å. If the molecular weight of rubber is 200,000, the length of the extended molecule would be 1.2 microns. If thermal vibration of its segments contracts it to one eighth of its full length, as is thought to be normal, this agrees with the above data and the fact that it clogs pores of 5,000 to 30,000 Å more or less rapidly.[18]

Light scattering has indicated a length of 3400 Å. Viscosity has given values for molecular or particle weight between 500 and 1,500,000. A smoked sheet gave a mean value of 3,000,000, but in all these cases the molecules are a mixture of very different sizes.

The Swelling of Rubber

Unvulcanized masticated rubber dissolves in many organic solvents such as kerosene, benzene, chloroform, and carbon disulfide (silver soaps dissolve in the same solvents) to form highly viscous, colloidal sols. In completely dry benzene it exhibits no electrokinetic effects. The benzene solutions containing moisture, however, exhibit typical electrophoretic effects and influences of added electrolytes. The viscosity is so great that a 0.5% solution has the viscosity of a 75% emulsion according to Hatschek's formula, Chapter 2. Unmasticated rubber and vulcanized rubber both swell without dissolving except as already described.

When rubber is exposed to nearly saturated vapors of suitable solvents, it takes up about two thirds of its weight of solvent, almost independent of previous treatment and vulcanization. The swelling of rubber has been very extensively investigated.

The swelling power of liquids for rubber lies in the following decreasing series: carbon tetrachloride, chloroform, tetrachloroethane, acetylene chloride, thiophene, toluene, benzene, cymene, cumene,

ethylene chloride, ether.[19] The swelling pressure is proportional to the cube of the concentration, the velocity of swelling being mono-molecular.

Synthetic Rubbers [20]

The extreme price fluctuations of natural rubber, 3¢ to $3 per pound, threatened world shortages of this indispensable material, shortcomings for particular applications, and, more especially, the desire of various civilized nations to have practically unlimited supplies have combined to make the production of synthetic rubbers one of the major industrial efforts of modern times.

The polyisoprene of natural rubber has never been duplicated by synthesis although attempts have been made for the last seventy years. It is an interesting fact that all artificial polymers tend to have the *trans* form and not the *cis* form, the latter being characteristic of natural rubber.

Before World War II the Bureau of Standards listed 29 kinds of "Synthetic Rubbers," several of which are no longer produced, and a number have since come into production. The following is a list of some of these prewar materials:

<div align="center">Synthetic Rubbers Manufactured Commercially</div>

Buna S (Germany)	Butadiene-styrene
Buna N (Germany); Perbunan (U.S.)	Butadiene-acrylic nitrile
Buna NN (Germany); Perbunan extra (U.S.)	Butadiene-acrylic nitrile
Neoprene (U.S.); Sovprene (U.S.S.R.)	Polychloroprene
Vistanex (U.S.); Oppanol (Germany)	Polyisobutylene
Resistoflex (U.S. and Germany)	Polyvinyl alcohol
Thiokol (U.S.); Perduren (Germany)	Polyalkylene sulfides
Koroseal (U.S.)	Plasticized polyvinyl chloride
Flamenol (U.S.)	Plasticized polyvinyl chloride
Knightware (U.S.)	Plasticized polyvinyl chloride
Igelite (Germany)	Plasticized polyvinyl chloride
Mepolam (Germany)	Plasticized polyvinyl chloride
AXF	Polyxylenes

It is noted that many of these are not rubber and not even hydrocarbons, but they are rubberlike in strength, stretching, and elasticity. They are more properly called elastomers.[21] It is clear that structure and molecular size play as great a role as does the chemical nature.

The American war effort was directed toward mass production, chiefly of Buna S, developed in Germany, which was a co-polymer of 6 molecules of butadiene, $CH_2=CH-CH=CH_2$, to one of styrene,

known as GR–S or Government Rubber Synthetic. Buna S is made up from the following: Bu stands for butadiene; Na, for natrium, the German and Latin name of sodium, the original catalyst; S, for styrene.

Polymerization may be carried out in a nonaqueous phase using a liquid monomer itself together with a catalyst, of which benzoyl peroxide is an example. It is of much greater interest for our purpose that polymerization is now usually conducted in aqueous emulsions which permit much closer control and reproducibility. The monomers are emulsified in soap solution, generally a tallow soap, but fatty acids of more than one double bond must be avoided because these retard the reaction. A number of other detergents may be substituted for the soap. There is added some strong oxidizing agent to promote reaction, such as persulfate or ferricyanide, and an agent, such as mercaptan, is added to minimize branching and cross-linking which would have the effects described above. Polymerization is initiated in the monomers solubilized in the lamellar McBain micelles from which the polymer is then extruded because of its high molecular weight. The soap is gradually exhausted in forming a layer around the polymer particles, thus producing a latex, and the final stages of the reaction proceed in these particles. Polymerization can occur wherever the soap molecules are arranged in ordered array, even, for example, at the interface between monomer droplets and water, if the interface is covered with the insoluble sodium stearate.

Every form of synthetic rubber has some advantages and some disadvantages in comparison with natural rubber. Polyisobutylene or Vistanex loses its only double bond during polymerization and is therefore resistant to oxidation, but it cannot be vulcanized. However, if a small proportion of butadiene is co-polymerized with it, the number of double bonds so introduced is sufficient for vulcanization, while it retains its stability. The result is butyl rubber, which has superseded all other forms of rubber for inner tubes of automobile tires because of its impermeability to gases. Butadiene itself, $CH_2{=}CH{-}CH{=}CH_2$, polymerizes to Buna SS or Buna 115, but the absence of side groups permits too ready crystallization and makes the product too stiff. Chloroprene, or chlorinated butadiene, $CH_2{=}CCl{-}CH{=}CH_2$, with the chlorine atom replacing the methyl group of isoprene, polymerizes to Neoprene, which is outstanding in resistance to gasoline and organic solvents and is therefore used for rubber hose carrying such products. It is also resistant to mustard gas and to Lewisite.

The rubberlike thiokols are made from sodium tetrasulfide with combinations of ethylene dichloride or dichloroethyl ether. They form

a useful latex. They are also very resistant to solvents and to deterioration but lack the outstanding physical characteristics of rubber. They yield a fiber diagram with x rays with axial periodicity of 5.7 Å for one form of fiber and 4.32 Å for the best formed fiber.[22] Cross-linking, analogous to vulcanization, can be effected by use of zinc oxide.
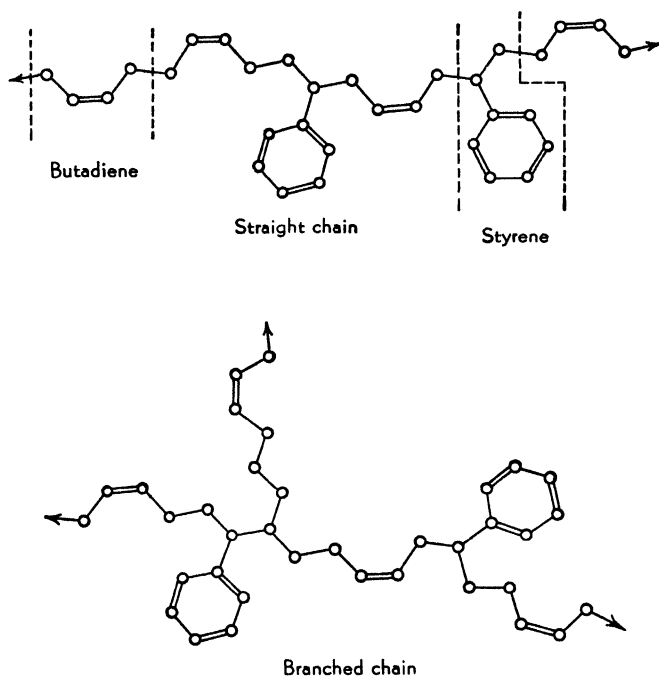
FIGURE 23-8 Branched and unbranched chain of Buna S. Taken from Treloar, *Nature*, **155**, 443 (1945).

The properties of all these synthetics are greatly dependent upon the details of their preparation as well as their subsequent processing, and they are all subject to continuous improvement. For example, Buna S has much better properties the lower the temperature of preparation, and I. M. Kolthoff announced in 1948 that by using an organic peroxide as catalyst it may be prepared in a reasonable time at temperatures as low as 0° C.

The particles of a GR–S latex are nearly one hundred times smaller than those of natural rubber — comparable with fine carbon black, as is readily shown by the electron microscope. The subsequent treatment of the synthetic latexes usually includes coagulation by destroying the emulsifying agent with acid or offsetting it with charges of opposite sign.

The individual molecules of a co-polymer, such as GR–S synthetic rubber, are all different in size and in the arrangement of the constituent groups. Flory [23] has stated, "It is unlikely that there are two identical polymer molecules in the many hundreds of thousands of this produced so far. The statement would hold true even if the amount of GR–S made equalled the earth in size." Crystallization of such irregular units is not possible. The contrast between an unbranched and a branched chain of Buna S is shown in Figure 23·8.

Neoprene varies in molecular weight, M, between 20,000 and 1,000,000 [24] by osmometry; the intrinsic viscosity obeys the formula

$$[\eta] = 1.46 \times 10^{-4} \, M^{0.73}$$

Other rubbery polymers are mentioned in Chapter 25.

Uses of Rubbers [25]

These colloidal materials have become indispensable in modern civilization. Some of the fields are transportation, including belting, de-icers, self-sealing tanks, compressed rubber, elastic springs, and shock absorbers — 2500 items used by the Army Air Corps. Then come surgical supplies, adhesive tape, cements, clothing of many kinds, and very numerous applications [26] in industrial equipment and in the textile industry.[27]

REFERENCES

Natural and Synthetic Rubber

BARRON, H., *Modern Rubber Chemistry*, 2nd ed., Hutchinson & Co., Ltd., New York and London (1947)

DAVIS, C. C. and BLAKE, J. T., editors, *Chemistry and Technology of Rubber*, Reinhold Publishing Corp., New York (1937)

DAWSON, T. R. and PORRITT, B. D., *Rubber, Physical and Chemical Properties*, The Research Association of British Rubber Manufacturers, Croyden, England (1935). This work assembles all the quantitative data on rubber up to 1935.

DUNBROOK, R. F. and MORRIS, V. N., editors, *The Science of Rubber*, Reinhold Publishing Corp., New York (1934)

FREUNDLICH, H., *The Chemistry of Rubber*, Methuen & Co., Ltd., London (1925)

LEWIS, W. K., SQUIRES, L. and BROUGHTON, G., *Industrial Chemistry of Colloidal and Amorphous Materials*, Macmillan Co., New York (1942), pp. 399–448

MARK, H. and WHITBY, G. S., *Scientific Progress in the Field of Rubber and Synthetic Elastomers*, Advances in Colloid Science, Vol. 2, Interscience Publishers, New York (1946)

MEYER, K. H., *Natural and Synthetic High Polymers*, Interscience Publishers, New York (1942), pp. 119–192

STAMBERGER, PAUL, *The Colloid Chemistry of Rubbers*, Oxford Univ. Press, London (1929)

Text References

1. TRISTRAM, G. R., *Biochem. J.*, **34**, 301 (1940)
2. MOYER, L. S., *Amer. J. Bot.*, **21**, 293 (1934)
3. HOMANS, L. N. S. and VAN GILS, G. E., *Medd. Nederland. Ind. Inst. v. Rubber*, Buitenzorg, Java, No. 67 (1948/9)
4. VAN GILS, G. E., *Trans. Inst. Rubber Ind.*, **33**, 74 (1947)
5. CLARK, F. E., *Chem. & Eng. News*, **26**, 926 (1948)
6. JONES, E. P., *Ind. Eng. Chem.*, **40**, 864 (1948)
7. GEHMAN, S. D., *Chem. Rev.*, **26**, 203 (1940)
8. KATZ, J. R., *Chem. Z.*, **49**, 353 (1925); *Naturwissen.*, **13**, 410 (1925)
9. HAUSER, E. A. and V. SUSICH, G., *Kautschuk.*, **7**, 120 (1931)
10. WALL, F. T., *J. Chem. Phys.*, **10**, 485 (1942)
11. KUBO, R., *J. Coll. Sci.*, **2**, 527 (1947)
12. ROBERTS, T. C., *Nature*, **141**, 834 (1938)
13. TRAPEZNIKOV, A. A., *Comp. rend. (Doklady) Acad. Sci. USSR*, **22**, 83 (1939)
14. GEHMAN, S. D. and FIELD, J. E., *Ind. Eng. Chem.*, **30**, 1031 (1938)
15. REHBINDER, P., AB, G. A. and WEILER, S. J., *Comp. rend. (Doklady) Acad. Sci. USSR*, **31**, 444 (1941)
16. STAMBERGER, P., *Kolloid-Z.*, **42**, 295 (1927)
17. PASSYNSKY, A. G. and GATOVSKAYA, T. V., *Nature*, **157**, 518 (1946)
18. MEYER, K. H. and JEANNERAT, J., *Helv. Chim. Acta*, **22**, 19 (1939)
19. POSNJAK, E., *Kolloidchem. Beih.*, **3**, 417 (1912)
20. WEBER, K. H. in POWETS, P. O., *Synthetic Resins and Rubbers*, J. Wiley & Sons, Inc., New York (1943), pp. 174–230
21. FISHER, H. L., *Ind. Eng. Chem.*, **31**, 941 (1939)
22. KATZ, J. R. and FULLER, C. S., Am. Chem. Soc. Meeting, April, 1935
23. FLORY, P. L., *Chem. Eng. News*, **25**, 1503 (1947)
24. NICHOLS, J. B. and MIGHTON, C. J., Chicago A.C.S. meeting, 1948
25. *Rubber in Engineering*, Prepared under the Direction of the Controller of Chemical Research of the Ministry of Supply (Great Britain), Chemical Publishing Co., Brooklyn (1946)
26. WOOD, L. A., Circular 427 of the National Bureau of Standards, 1940
27. HAUSER, E. A., *Rayon Textile Monthly* (Sept., 1937)

THE PROTEINS SILK, WOOL, KERATIN, COLLAGEN,

AND THE GLOBULAR PROTEINS

As was stated in Chapter 18, the proteins are polypeptides derived from the substituted levo-α-amino acids. One of the simplest is the *fibroin* of silk, which serves as an example of the elongated *trans*-polypeptide chain. A second large group of fibrous proteins is that of the *keratins* and the *muscle proteins* such as myosin (and fibrinogen).[1] The keratins include the hair and fur of all animals, feathers, nails, claws, quills, whalebone, horn, spines, etc. As a matter of general interest, the appearance of human hair under the microscope is shown in Figure 24·1. The keratins are very different from silk in that in the natural state they are folded and often cross-linked. However, when stretched they are like silk in structure. A third group comprises the *collagens*, which are *cis*-polypeptides, permanently folded, and which cannot readily be stretched as a result of their proline and hydroxy groups. These three groups are essentially fibers.[2] The fourth group is that of the *globular* or *corpuscular proteins*, which are completely folded, such as the naturally occurring particles of egg albumin and hemoglobin. It is the globule or particle that occupies particular points or rather spaces in the crystal space lattice.

Polypeptides of very high molecular weight have been synthesized by the reactions due to Leuchs (1906), Go and Tani, Woodward and Schramm, and others.[3] Bergmann [4] developed methods of preparation and analysis in which the α-acids are added or detached *seriatim* so that their positions can be identified.

Miss Wrinch has pointed out very emphatically the contrast between proteins and other high polymeric substances.[5] "Proteins, in
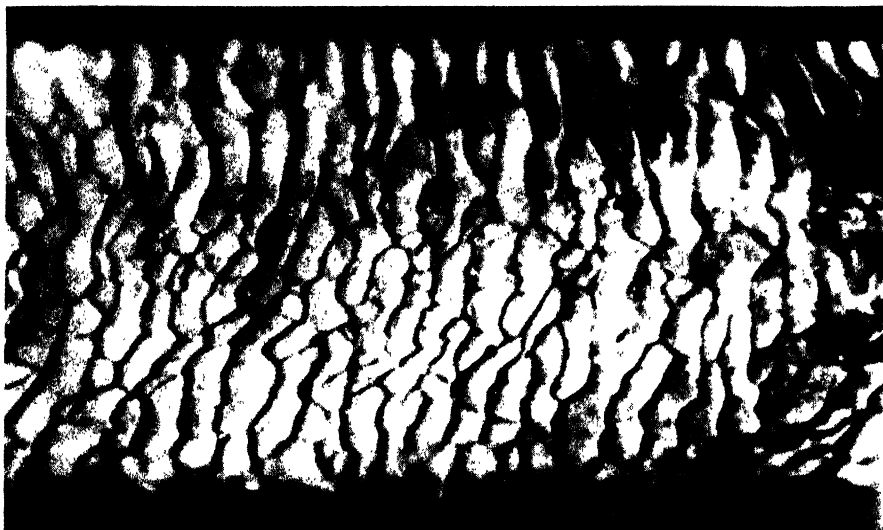
FIGURE 24·1 Photomicrograph of human hair (original size about 100 microns). Courtesy of J. Manby.

general, crystallize — and indeed maintain their existence — only with the aid of foreign molecules or ions, notably water. . . . Proteins, with their tendency to denature, are intrinsically unstable, a characteristic setting them apart from most — possibly all — other substances."

Silk

Natural silk consists of a water-insoluble, mechanically resistant fibrous material known as fibroin, with an outer coating of 15–20% of sericin, and upon this a fraction of 1% of wax. The sericin is removed by boiling in soapy water. The silk fibroin consists to the extent of nearly half of glycine residues,

$$\text{NH} \quad \text{CO}$$
$$\diagdown \text{CH}_2 \diagdown \text{,}$$
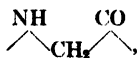
and over a quarter of alanine residues,

$$\text{NH} \quad \text{CO}$$
$$\diagdown \text{CH} \diagdown \text{,}$$
$$\overset{|}{\text{CH}_3}$$
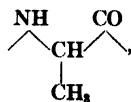
some tyrosine, with very minor amounts of other amino acids. These are combined to form fibers several hundred residues long. The repeat-

ing portion of the backbone of the polypeptide structure is 3.5 Å long, or a multiple thereof such as 7 Å or 10.5 Å, quite independent of the nature of the side chains or the constituent amino acids. The side spacings obtained by x rays are 4.3 and 4.6 Å. The x-ray fiber diagram, shown in Figure 24·2, is imperfect, showing that part of the silk is amorphous and part crystalline. The tyrosine portions, with a relatively bulky p-cresol side group, are too large to be included in the crystalline regions of the fibroin.

The tensile strength and x-ray pattern of wet silk are the same as for dry silk, showing that water does not enter into the crystalline portions, which accounts for the low water absorption of silk.

Von Weimarn [6] has shown that silk may be dissolved in concentrated solutions of sodium thiocyanate and precipitated again in the form of a glass, or it may be spun, stretched, and stroked until it is far
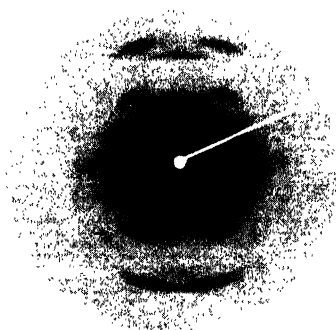
FIGURE 24·2 X-ray fiber photograph of silk. Taken from Senti, F. R., *Am. Dyestuff Reptr.*, **36**, 230 (1947).

stronger than natural silk but more brittle. From this structure it is evident that silk fibers possess great tensile strength but that, like cellulose, they are not very elastic. After being stretched beyond a certain point they will take a permanent set on account of the slippage of the fibers.

Silk fibroin is the first protein to be completely renatured after denaturation.[7] It is dissolved in aqueous cupri-ethylenediamine, neutralized, and dialyzed. The molecules are about 1300 Å long with a weight of 33,000 parallel to the fiber axis. Acidification to pH 3 causes immediate precipitation of the denatured protein. Mere evaporation gives a water-soluble, cellophane-like film that can be stretched when moistened to about three times the original length, whereupon it is

completely insoluble in water, strongly birefringent, and gives an x-ray diagram identical with that of the original fibroin, which in the following paragraphs we shall identify as the β-keratin or elongated polypeptide chain structure. Hence denaturation consists chiefly of the unfolding of a molecular chain into the completely extended form, but with fibroin the folded chain is not the α-keratin structure of wool, to be discussed below.

The isoelectric point of silk fibroin is pH 4.7. Donnan equilibria and base exchange phenomena are prominent in the processing of silk.

Wool and the Other Keratins

The keratins are insoluble in all solvents. They swell to a limited extent and are relatively resistant to the action of hydrolytic agents. Astbury and Street [8] were the first to elucidate their x-ray structure. They distinguished between an α-form unstretched and a β-form under tension, corresponding to Figures 24·3 and 24·4, respectively.

The first thing one notices about the x-ray photographs is how ill-defined they are, showing poor and only partial orientation. We look at once for the repeating unit of 3.5 Å of the stretched polypeptide chain as in silk but cannot find it in the unstretched or α-keratin. Instead the repeating unit is 5.1 or 10.2 Å. Thus wool or α-keratin is quite different from silk. Models of silk and α-keratin are contrasted in Figure 24·5. The most perfect examples of crystallized α-keratin
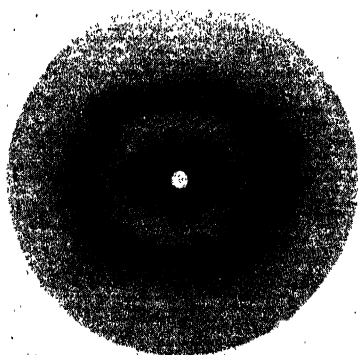
FIGURE 24·3 X-ray fiber photograph of alpha-keratin, *viz.*, English Cotswold wool, soap-scoured, zero per cent extension. Courtesy of W. T. Astbury.
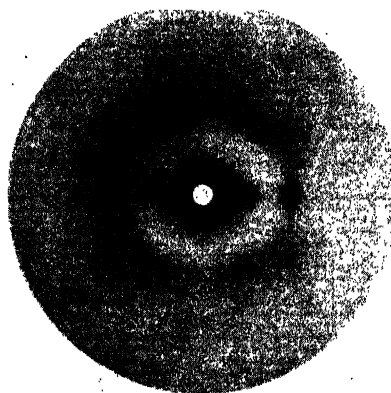
FIGURE 24·4 X-ray fiber photograph of beta-keratin, *viz.*, English Cotswold wool, soap-scoured, 80 per cent extension. Courtesy of W. T. Astbury.
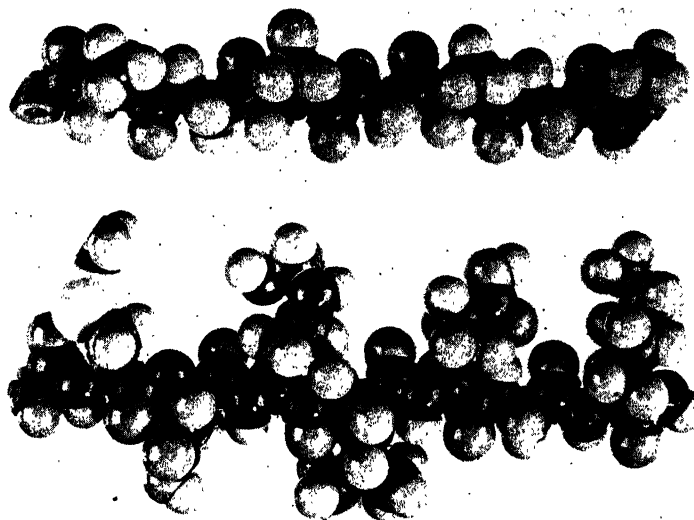
FIGURE 24·5 Photograph of scale models of polypeptide chains of a silk (above) and wool (below). Taken from Senti, *Am. Dyestuff Reptr.*, **36**, 230 (1947).

are provided by quills, as for instance those of the gull or porcupine. Figure 24·6 is the diagram showing the x-ray diffraction that can be identified in a photograph of porcupine quill.

Stretched or β-keratin does show the repeating unit of 3.4 Å of the stretched polypeptide chain, like silk. In the crystallized β-keratin the polypeptide chains lie parallel, like backbones of fish, 4.5 Å apart, but the side spacings have cross-linkages holding the chains 10 Å (7–11 Å) apart in the direction of the zigzag. The three spacings, 10.2, 4.5, and 3.4 Å are at right angles to each other. Hence, like cellulose, the properties are quite different in the three directions.

β-keratin is a denatured protein, but it is quite disorganized unless it has been stretched to crystallize it. Since denatured keratin is the form obtained by spreading on water, the thickness of a monolayer is nearly 10 Å. This received an intriguing confirmation with a film built up to several thousand layers by the Langmuir-Blodgett technique where the over-all thickness could be measured with ordinary calipers.

Ambrose and Hanby [9] have examined a synthetic polypeptide derived from a substituted glutamate using polarized infrared radiation. They adduce evidence for folding of the peptide chains as follows, where the dots represent hydrogen bonds within the folded chain,
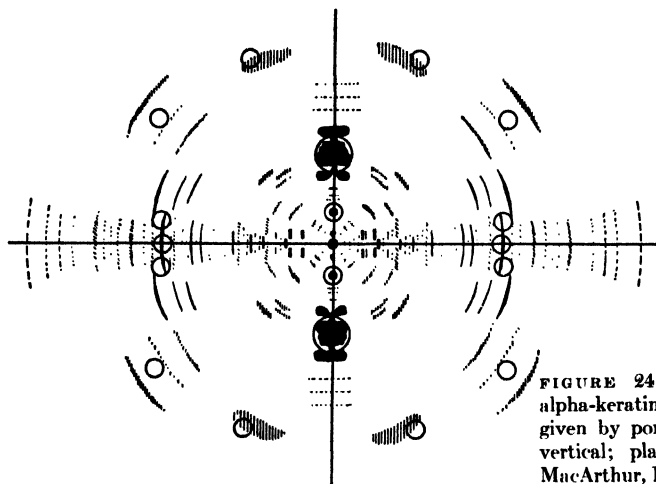
FIGURE 24·6 Composite chart of
alpha-keratin x-ray diffraction pattern
given by porcupine quill. Fiber axis
vertical; plane plate. Taken from
MacArthur, I., *Nature*, **152**, 38 (1943).

while the side chains and seven membered rings both lie approximately
in the plane of the ribbon.

The transformation from α- to β-keratin involves some sort of
unfolding or straightening of the polypeptide chain, with a doubling
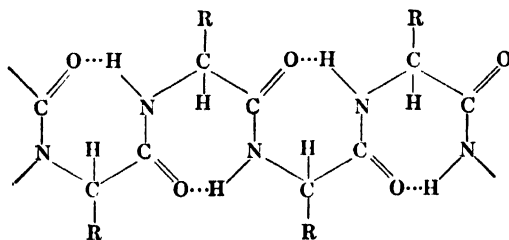
of its length. The explanation put forward by Astbury and Street [10]
and which was current for many years thereafter is represented in the
following diagram:
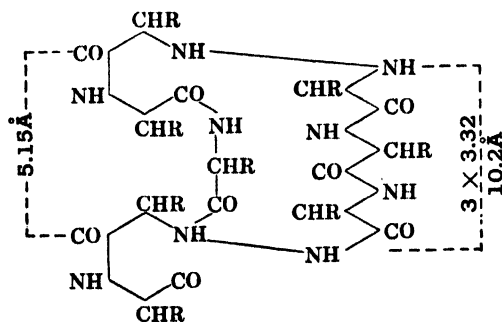
This particular explanation had to be abandoned because in the hexagonally folded form suggested for α-keratin there was simply not room for the interatomic spacings and particularly for the side groups which, after all, often amount to more than half the total amount of the protein.

Hence, Astbury and Bell [11] proposed instead that the polypeptide chain lies in a plane of folds alternately one way and then the other, each following three sides of a square of 5.1 Å per side, so that when the chain is pulled out straight it doubles the repeating length to 10.2 Å. The opening of each fold involves the breakdown of 3 hydrogen bonds between CO and NH groups. Then the side groups are packed closely together above and below this plane. For supercontraction, mentioned below, the folds are merely made longer. Proteins are then supposed to consist of a number of such sheets or layers superimposed to form micelles. The number suggested for native egg albumin is 4 such sheets.

Wool in its normal state, α-keratin, consists of chains that are not merely coiled or folded but are cross-linked by salt linkages between carboxyl and amino groups on the side chains, by hydrogen bonds, and chemically by primary valences through the cystine linkage, $-CH_2-S-S-CH_2-$. The cystine link can be broken without disintegration of the fiber substance. It is this breakdown which is fundamental to felting, and wool therefore owes this unique property to its sulfur content. The cystine linkage is broken most readily by alkaline solutions, the salt linkages by acid solutions. The worm or larva of the clothes moth secretes a powerful reducing agent which attacks the sulfur cross-linkages; then the exposed chains can be attacked by proteolytic enzymes.

The isoelectric point of wool has been reported as between pH 4 and pH 8, but the isoionic point of the wool itself has been found to be between 6.0 and 6.2.[12] The cross-linkages hinder extension and exert a contracting influence when the stretching force is removed. It is the tendency to contract to the folded α-form, a tendency which is reinforced by the cross-linkages, that makes wool porous and gives woolen garments their unique heat insulating property.

The behavior of wool is greatly dependent upon moisture. In water, wool fiber becomes 18% greater in diameter but only about 1% longer, the water being chiefly on the outer surface of the crystallites which appear to be not more than about 200 Å thick. In saturated vapor or in water, wool takes up about 33% by weight, but at 63.4% relative humidity it takes up 14% of water vapor.

This reduces the insulating power by 10^{12}-fold and reduces the rigidity 15-fold.

As regards stretching and contracting, "what heat is to rubber, water and especially hot water is to wool." Dry wool cannot be stretched more than 20%. Immersion in cold water enables it to stretch up to 70%. In hot water or steam wool is readily stretched 100% to β-keratin. Under water this is perfectly reversible. Wool is elastic because of its tendency to contract to the normal α-keratin form. β-Keratin cannot shrink when dry.

Wool if stretched dry will always shrink if afterwards wetted. Wool that has been stretched not more than 30%, if then placed in water without tension, will recover practically its original length, but if it is stretched in hot water or steam to 100% and then dried, it remains as extended β-keratin, as in the so-called permanent wave of the beauticians.

When wool is stretched in steam and then released in the steam, it contracts to a length about one third shorter than the original α-keratin. This is called supercontraction. The white of a boiled egg is in this state. The supercontracted form gives a much more amorphous looking x-ray pattern, termed "cross-β."

The protein from epidermis exhibits the three forms, α-, β-, and supercontracted keratin, which are readily converted into each other. Supercontracted epidermis is reconverted to the α-form by means of saturated urea.[13]

The Australian investigators Mercer, Rees, and Farrant [14] have significantly extended the knowledge of wool fibers and the keratins therein. The wool fiber is not hollow but is a bundle of long cells held together by a matrix of disordered keratin in which denser fibrils are embedded. The matrix flows under stress like a viscous liquid and readily changes from sol to gel to sol. Its molecules, as interpreted from electron microscopy, are about 110 Å in diameter. The fibrils are composed of protofibrils which in turn are composed of linear aggregates of approximately isometric particles or molecules of molecular weight 360,000, 110 Å in diameter, which indicates a closer association with the globular proteins, next to be discussed, than with the Astbury conceptions of keratin chains. These molecular diameters correspond to the feather keratin fiber period actually obtained by low angle x-ray scattering and to half the very long x-ray spacing, 198 Å, of the porcupine quill, which is otherwise unexplained. The Australian investigators suggest that fibrous α-keratins consist of strings of globular molecules rather than the folded structures described above.
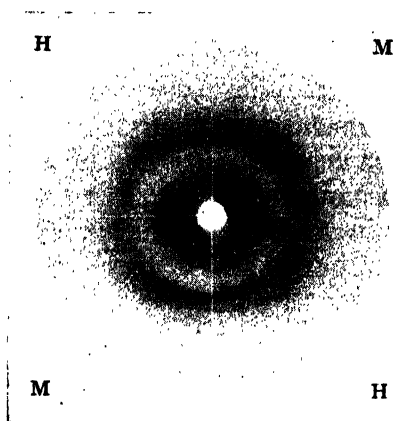
FIGURE 24·7 X-ray sector comparison photograph of alpha-horn (H) and alpha-myosin (M). Courtesy of W. T. Astbury.
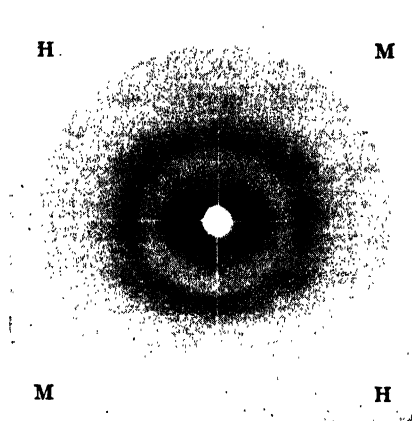
FIGURE 24·8 X-ray sector comparison photograph of beta-horn (H) and beta-myosin (M). Courtesy of W. T. Astbury.

A cuticular scale cell wall has embedded within it a more resistant component not digested by trypsin which attacks the remaining keratin.

As long ago as 1938, Harrison [15] criticized the overemphasis on the fraction of the keratin which x rays show to be in the crystalline state and instead emphasized that sorption and desorption of water and ions and chemical reactions play an important part in all the swelling and deswelling phenomena associated with keratins and only in that way do such reactions play a part in supercontraction, permanent set, and in other properties of keratins.

Muscle or myosin is normally in the supercontracted form, as deduced by Astbury from x-ray evidence. It is significant that it is relatively free from sulfur and cannot therefore be strongly cross-linked. In contrast, hair, with its 5% of sulfur, may be likened to vulcanized myosin. Photographs of myosin and horn in the α- and β-forms are given in Figures 24·7 and 24·8. Astbury's views have been summarized by him in the Croonian lecture of 1945, printed in 1947.[16]

During the last few years Szent-Györgyi and his collaborators [17] have developed a completely new contribution to the problem of the muscle proteins. They find that the myosin of previous workers as referred to above was not a protein of definite composition but is a combination of two proteins, each of which can be prepared in crystalline condition. For one they retained the name *myosin* and to the other they gave the name *actin*. The complex of the two, the old myosin, they called *actomyosin*. In the *Journal of Colloid Science* Szent-Györgyi [18]

has given a clear, concise, and simple account of the colloid behavior of myosin and of actin, as well as of actomyosin.

The shift of the isoelectric point of proteins by salts is greatly exaggerated in myosin. When dissolved in potassium chloride, its isoelectric point is at pH 5.4, but in calcium chloride it is at a pH over 8.7.[19] Actin is another lyophilic colloid with isoelectric point at pH 5.7. Like myosin it can exist in either fibrous or crystalline globular form.

Actomyosin swells enormously in pure water, but not in the presence of even very dilute salts. When precipitated it contains from 95 to 98% water. High concentrations of salt dissolve it.

Szent-Györgyi pays but little attention to the x-ray evidence but instead stresses the effect of the loss of the small amount of bound hydrate water of the myosin only (not the large enmeshed amount) in causing a one-sided contraction of the actomyosin micelle. His mechanical model is given in Figure 24-9, in which one side, a slightly stretched piece of rubber representing the hydrated myosin, has attached on its side the string of the smaller globular particles of actin. On releasing the rubber (dehydration of the myosin), its slight contraction causes the model to curl up and contract greatly.

It must be remembered that four proteins are usually recognized as being components of muscle. These are the myosin or actomyosin, amounting to only 39%, the water soluble myogen, the muscle globulin, and the insoluble stromoprotein. Contracted muscle shows an "amorphous" x-ray diagram.
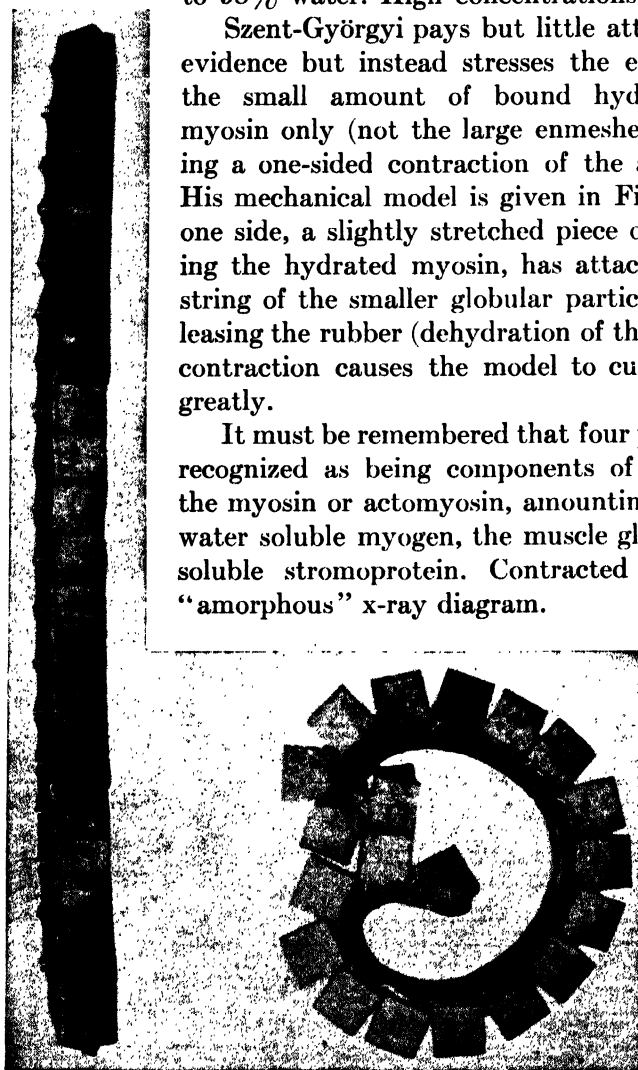
FIGURE 24-9 Mechanical model of actomyosin micelle, after Szent-Györgyi.

Globular or Corpuscular Proteins

Quite surprisingly, many soluble proteins can be obtained in the crystalline condition. This was queried for many years although some preparations looked like crystals, as is seen in Figure 24·10. Examination of the precipitated and dried proteins had yielded only an amorphous pattern, but Bernal and Crowfoot [20] discovered that if protein crystals are examined in their mother liquor without exposure to air, they give a very highly developed crystal pattern. Since then wet "globular" proteins have been prepared in all the crystal systems. Among the earlier ones to be studied were pepsin, insulin, edestin, excelsin, hemoglobin, the Bence-Jones protein of Figure 24·10, and other globulins and enzymes. The patterns resemble those of feather keratin so far as the arrangement of molecules is concerned. Insulin crystallizes in the rhombohedral system.[21] It is the only protein in which the unit cell contains only one particle. It has a weight of 39,700, which includes 5.35% water, leaving 37,600; but there is no crystallographic evidence that the contents of this unit cell constitute a single molecule. Like hemoglobin, there are present at least two different polypeptide chains.[22]

For β-lactoglobulin in the form of orthorhombic crystals wet with mother liquor, a molecular weight of 34,500 is deduced from x-ray examination.[23]

FIGURE 24·10 Crystal of Bence-Jones protein; magnification of original photograph about 225 ×. From Magnus-Levy, Meyer and Lotmar, *Nature*, 137, 616 (1936).

By far the most elaborate and careful study of any globular protein is that of Boyes-Watson, Davidson, and Perutz.[24] Horse methemoglobin crystallizes in the monoclinic space group C_2 with two "molecules" of dry weight 66,700 per unit cell. In addition, 52.4% of the unit cell volume is liquid of crystallization. These crystals can be swelled or shrunk, or the composition of the liquid may be changed by dissolving various ions in it. At the isoelectric point, pH 6.8, this may be done, or all ions may be removed without change in crystal dimensions.

Figure 24·11 shows diagrammatically the structure of this hydrous crystal. The cylinders are the methemoglobin molecules 34 Å high

FIGURE 24·11 Diagram showing packing of methemoglobin molecules in the crystal structure, showing layers of close-packed molecules separated by liquid. One unit cell is shown in the foreground on the right. Taken from Boyes-Watson, Davidson, and Perutz, Text Ref. No. 24.

and 57 Å wide, lying in close-packed layers. These layers themselves do not swell or shrink either in thickness or in area except on complete drying of the bound water between them, which is not available as solvent to diffusing ions. Between the layers of hemoglobin molecules is a layer of free water 15–25 Å thick, accurately spaced under any definite conditions, just as in the micelles of soap and detergent solutions or as in crystals of hydrated clay. The hemoglobin molecules themselves consist of two chemically and structurally identical halves, as is shown in Figure 24·12. An indication of how remarkably sharp and highly organized these crystals are is given in Figure 24·13.

It is evident that there is a wealth of information to be gathered from a detailed study of these hydrous systems. They exhibit successive levels of organization down to quite small blocks, with emphasis on spacings around 10 Å. However, the bewildering complexity may prohibit complete interpretation of the overabundant evidence.

Human carboxy hemoglobin has been crystallized [25] in quite a different class and space group with four molecules to the orthorhombic cell whose dimensions are 85.2, 77.0, and 86.0 Å when wet, and when dry, 70.4, 73.3, and 77.0 Å. The external form is that of bipyramidal prisms of almost square cross section. There are again many spacings in the neighborhood of 10 Å.

Many of the globular proteins are very soluble in water or salt solutions. Denaturation of the globular proteins results in the forma-

tion of β-keratin, the fibrous form. Similarly, spreading upon water yields monomolecular layers which are extremely insoluble.

Lundgren [26] discovered that by means of detergents it is possible to unfold the native globular proteins and then to draw the precipitated protein into elastic and highly doubly refracting fibers. This he first did with egg albumin, hog thyroglobulin, wheat glutenin, casein, zein, and blood albumin. Many properties of these solutions have been studied. His work is discussed in a series of subsequent papers.[27] This very interesting research has great commercial importance.

Reduced chicken feather keratin forms a soluble complex in aqueous sodium dodecyl benzene sulfonate. The molecular weights of the two kinds of dispersed keratin in this solution are 34,000 and 41,700. The solution is spun and the detergent removed. The tensile strength of the resulting fiber is up to 80,000 lb/sq in, that from egg 20,000–70,000, casein and soybean up to 10,000, as compared with 46,000–74,000 for silk and only 17,000–25,000 for wool.

In addition to this work from the Western Regional Laboratory, much has been done in the Eastern Regional Laboratory using high concentrations of organic solvents in place of dilute solutions of detergents.[28]
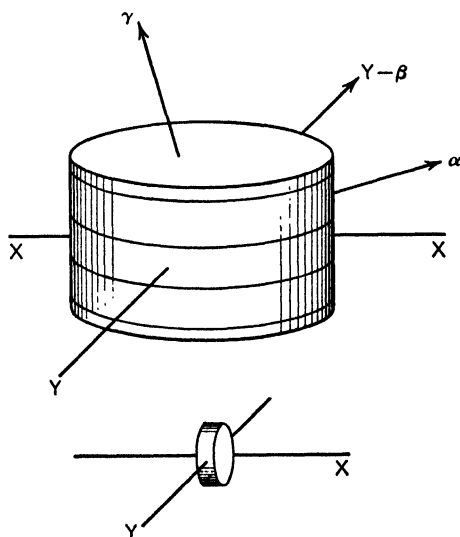
FIGURE 24·12 Diagrammatic model of hemoglobin molecule, showing its orientation with respect to the crystal axes. Y is the diad axis. The small disk underneath represents a hemo-group drawn on the same scale and in its correct orientation with respect to the crystal axes. The four lines on the cylinder surface indicate the positions of the concentrations of scattering matter deduced from the Fourier projections. The directions of the principal refractive indices are indicated by arrows. Taken from Boyes-Watson, Davidson, and Perutz, Text Ref. No. 24.
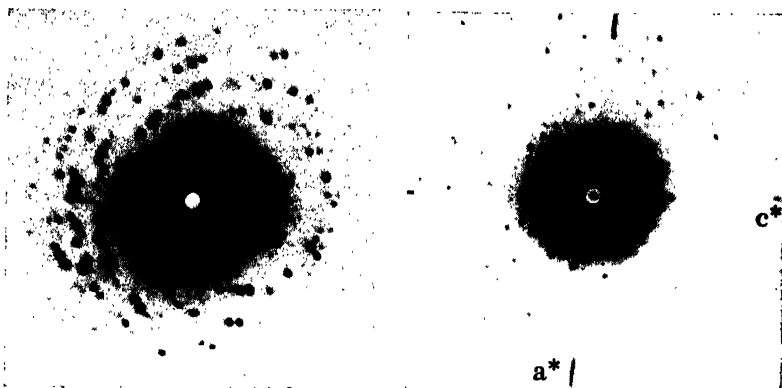
FIGURE 24·13 Oscillation x-ray photographs of horse methemoglobin crystals. Photograph at left is from the wet crystals taken on a flat film at 5 cm distance; oscillation about 110, beam parallel, 001 at 6.5–9°. At right is a 5° oscillation photograph of the same material in a different state of swelling. Film distance 6 cm; crystal oscillated about 100 and the beam parallel to 010 at one end of the oscillation. The dark halos near the center are due to air scattering. Their intensity is much exaggerated by the reproductions. Taken from Boyes-Watson, Davidson and Perutz, Text Ref. No. 24.

It is thus in general possible to convert a globular protein or a virus into insoluble fibers of β-keratin and to transform these into bristles or textiles. It follows then that, in theory, it is possible to convert even tobacco mosaic virus into textiles, "an overcoat from a disease."

The nature of the folding of the corpuscular or globular proteins has been much discussed, and it is usually considered that the folding is in layers, with probably four layers in native egg albumin. Lundgren's conception of the relations between the forms of such a protein is summarized in Figure 24·14.

The Collagens

The collagen or elastin group comprises the white connective tissue fibers, tendons, cartilage, the scales and fins of fishes, the ichthyocol of swim bladders, the byssus threads of bivalves, the so-called ovo-keratin of the egg capsule of the skate, etc., the cuticles of annelid worms, jellyfish, the filaments ejected by the sea cucumber, and so on. It is characterized by polypeptide chains which are held inextensibly in a configuration of the β-proteins, and the evidence to date is that this is the result of the preponderance of residues of the two amino acids proline and hydroxyproline, which have the effect of impressing a partial *cis* configuration on the chains.[29]

An electron micrograph of cowhide collagen is given in Figure 24·15. The first point that meets the eye is the banded structure at a uniform repeating distance of 640 Å.[30] Each of these "bands" contains six subunits unequally spaced. The diameter of a typical fibril is 800 Å. With special x-ray examination at very low angles, O. E. A. Bolduan
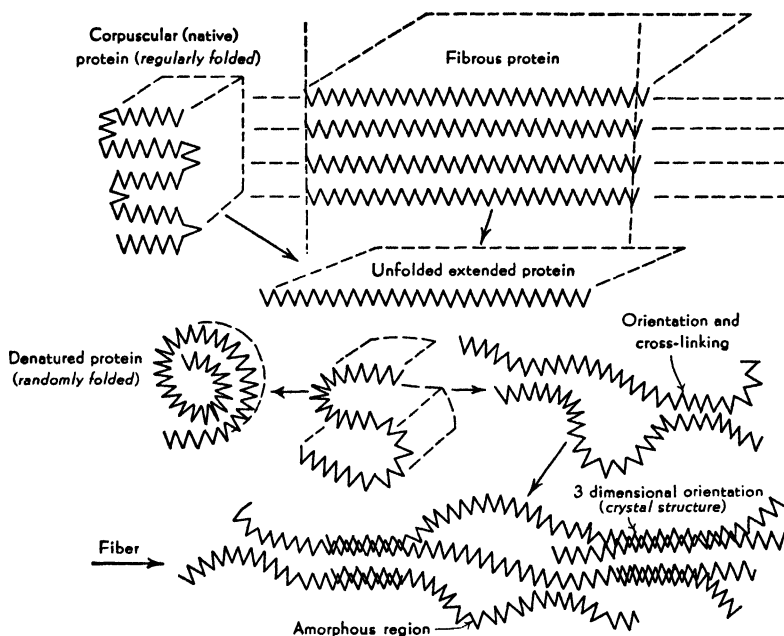
FIGURE 24·14 Lundgren's representation of structural transformations involved in fiber formation from corpuscular and fibrous proteins. In corpuscular proteins chains are restrained in regularly folded configurations; in fibrous proteins they are restrained in extended configuration. In fibers, portions of chains are restrained in extended configuration. Taken from *Textile Res. J.*, 15, 338 (1945).

has obtained no less than 33 orders of this long spacing. Long spacings have been observed for many other naturally occurring high polymers such as cellulose, silk, and keratin, but their significance has scarcely been discussed.

The long spacing of 640 Å has been found [31] to vary with the relative humidity of the atmosphere from 628 Å at 2% to 672 Å at 100% relative humidity. This has been interpreted as being due to water entering between the collagen molecules as in the structure of hemoglobin crystals already referred to.

The repeating distance along the polypeptide chain is only 8.4 Å, which is unaffected by stretching or relaxing the fibers. Evidently much of the collagen is not in the oriented or crystalline portions.

About one third of the residues in collagen are glycine; somewhat fewer are proline or hydroxyproline, with only very minor quantities of the remaining amino acid residues. The x-ray side spacings are 4.65 and 9.5 Å.

An important property of the collagen group is that of thermal contraction. In hot water and in certain swelling solutions the collagen fibers contract very rapidly to as little as one fifth of their ordinary length and then show long rubberlike elasticity. With x rays this super-contracted collagen shows only amorphous halos. In all media which have no hydrolytic action, collagen is insoluble and swells only to a limited extent. The molecular weight of native collagen cannot therefore be determined.

Collagen fibers exhibit fraying at the ends, and the action of trypsin and bacteria is limited to the damaged or cut ends of the fibers. The chemistry of tanning is the chemistry of turning the hydrated and putrescible collagen into the dehydrated material, collagen tannate, with its active centers suppressed and the peptide links protected from the liability of putrefactive attack. In hides and skins the tanner has an excellent physical material possessing firmness and flexibility but also a chemical instability and liability toward bacterial attack until it has been tanned to make the exposed sensitive surface hydrophobic.[32] Tanning reagents are those which precipitate gelatin in insoluble form. Vegetable or chrome tanning causes a great loss of definition in the x-ray spacing due to the side chain.[33] Formaldehyde and syntan tanning do not affect the x-ray pattern [34] but provide strong permanent cross-linkages. Both chromium salts and tannin, being polyvalent, link neighboring chains together, but their action is less strong and can be reversed.

The shearing action of a cutting tool alters the character of collagen, making it take on the properties of gelatin, through a change of molecular structure rather than chemical nature. An outstanding property of collagen is its swelling in hot water with the gradual formation of the water-soluble product gelatin through hydrolysis of main chain peptide linkages. Prolonged heating of gelatin in water continues to break the chain by hydrolysis and hence to reduce the viscosity. The location and distribution of the water molecules bound by hydrogen bridges to the gelatin chain has been ascertained by x-ray analysis.[35] Gelatin is digested by trypsin.

Hot, dilute gelatin solutions display no marked structural viscosity, but through association of the gelatin molecules on cooling, larger
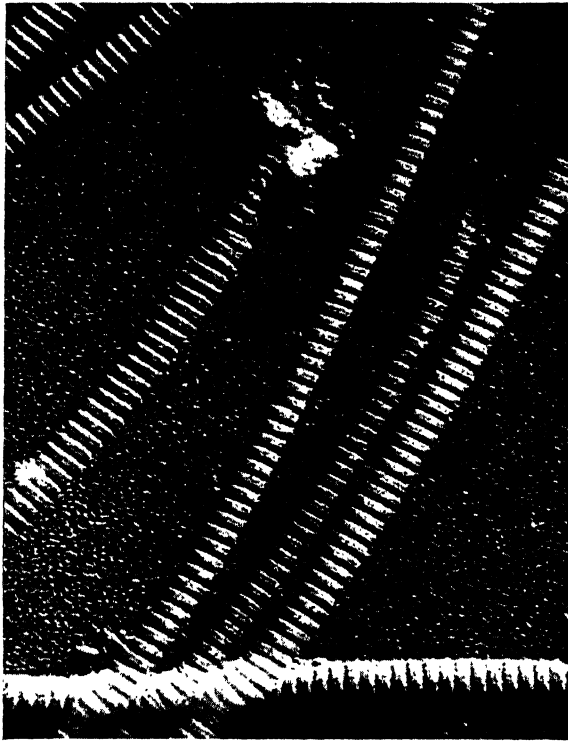
FIGURE 24-15 Electron micrograph of cowhide collagen
fibrils shadowed with gold. Magnification of original photo-
graph 33,500 ×. Taken from Senti, *Am. Dyestuff Reptr.*,
36, 230 (1947).

aggregates and micelles are formed, and a brush heap jelly structure is
developed. This illustrates the importance of association even with
high polymers. It may be recalled that the association of the colloidal
particles in a gelatin sol to form jelly does not measurably influence
the electrical conductivity, which shows what a loose connection be-
tween primary particles is required to produce a jelly structure. The
rigidity of dilute gelatin jellies is proportional to the number of useful
cross-links and the square of the concentration.

The swelling of gelatin jellies was correlated with the Donnan
equilibrium by Procter and Wilson.[36]

Gelatin jellies containing aqueous alcohol or glycerin can be
stretched several hundred per cent; the x-ray diagram characteristic
of collagen then makes its appearance, and the gelatin becomes solid
in the direction of stretching but splits readily at right angles to
this.[37]

Nucleoproteins and Viruses

Proteins have often been discussed without mention of minor non-protein constituents, which play a major role in their behavior, such as the nucleotides, carbohydrates, and lipids.

> The x-ray study of the polynucleotides, like the chemistry of these all-important structures, is as yet practically in its infancy. An encouraging start has been made with sodium thymonucleate, the diffraction pattern of which indicates a columnar piling of flattish nucleotides spaced at almost exactly the same distance as the separation of successive side-chains in an extended polypeptide,[38] and similar features are suggested by preliminary photographs of nucleic acids of the ribose type; but otherwise all this great world remains to be conquered. The proffered reward is tremendous, for, whatever else is in doubt, it seems clear enough now that the innermost secrets of biosynthesis rest above all with the proteins and the nucleic acids.[39]

The viruses and chromosomes are nucleoproteins. Stanley[40] has made the most important contributions to the study of the virus proteins by isolating and purifying them and finding that they can be prepared in a fully crystalline state. Wyckoff and Corey[41] have shown that tobacco mosaic virus gives the x-ray pattern of a true crystal, composed of large molecules and having many sharp reflections between 3 and 80 Å. The particles of tobacco mosaic virus have "a molecular weight of at least 46,000,000, a length of at least 4300 Å, and an effective diameter of about 120 Å."[42] The dimensions have been both examined directly in the electron microscope and deduced from such measurements as ultracentrifugation, viscosity, and light scattering. Both aggregation and subdivision of these particles are prominent. They form long fibers visible in the microscope which are easily broken up by mechanical means.[43]

More recently Stanley and Anderson[44] decided that the true length was 2800 Å. Frampton[45] has discussed the variation in their photograph from 370 Å to 3400 Å, as well as those of others. In a recent photograph the length of particle in the same specimen ranges from 2700 to 3500 Å, the middle values preponderating. Oster, Doty, and Zimm[46] obtained a reasonably monodisperse specimen whose lengths in the ultramicroscope averaged by three methods were 2620, 2690, and 2710 Å. With the same specimen they deduced from the viscosity 2600 Å, and from light scattering 2700 Å, and a molecular weight of 40 million, ± 5%.

Bernal and Fankuchen[47] have examined by x rays tobacco mosaic virus and other varieties of virus proteins. Of greatest interest were

their observations of the mutual arrangement of the virus particles in the anisotropic phase.

The long molecules of tobacco mosaic virus can be easily oriented; and oriented preparations ranging in concentration from dry gels to aqueous solutions as weak as 2% have been studied. The x-ray patterns indicate that the same sort of molecular packing exists over the entire range of concentrations. The molecules are parallel, and in cross-section the arrangement is hexagonal. The x-ray data enable the direct computation of interparticle distances to be made. For specimens of different water content it was found that the interparticle distance was a simple function of the concentration.

The Reproduction of Proteins

It has long been recognized that proteins with their highly specific properties must result from some template method of production from pre-existing proteins. The globular proteins show an extreme degree of specificity. Thus hemoglobin in each different type of animal seems to differ in spectrum and in detail of its behavior with oxygen. Immunochemistry in general draws the same picture of globular proteins as highly organized structures with extremely delicate specificities, as evidenced by antigen-antibody reactions.[48]

Mirsky and Pauling [49] attributed the characteristic specific properties of proteins to a folding of the polypeptide chain into a uniquely defined configuration held chiefly by hydrogen bonds. Antibodies have the chemical characteristics of serum globulins. Apparently unlimited numbers may be produced, possibly by the detail of folding. Films of protein only 9 Å thick are still capable of reacting with antibodies,[50] but the activity of antibodies is destroyed by such unfolding of the molecule.[51]

A detailed although diagrammatic account of the formation, structure, and behavior of antibodies and their antigens has been given by Pauling.[52] This provides the framework theory of serological precipitates. However, some of the evidence and conclusions have been questioned.[53]

REFERENCES

The Proteins Silk, Wool, Keratin, Collagen, and the Globular Proteins

MEYER, K. H., *Natural and Synthetic High Polymers*, Vol. IV, Interscience Publishers, Inc., New York (1942), pp. 423–564

Text References

1. BAILEY, K., ASTBURY, W. T. and RUDALL, K. M., *Nature*, 151, 716 (1943)
2. *Fibrous Proteins*, Proceedings of a Symposium held at the University of Leeds, May, 1946, published for The Society of Dyers and Colourists, by Chorley and Pickersgill, Ltd., Leeds, England
3. HANBY, W. E., WALEY, S. G. and WATSON, J., *Nature*, 161, 132 (1948)
4. BERGMANN, M., *Science*, 79, 439 (1934)
5. WRINCH, D., *Science*, 107, 445 (1948)
6. VON WEIMARN, P. P., in Vol. IV, J. ALEXANDER'S *Colloid Chemistry*, The Chemical Catalog Co., New York (1932), p. 363
7. COLEMAN, D. and HOWITT, F. O., *Nature*, 155, 78 (1945)
8. ASTBURY, W. T. and STREET, A., *Phil. Trans. Roy. Soc.*, 230A, 75–101 (1932)
9. AMBROSE and HANBY, *Nature*, 163, 484 (1949)
10. ASTBURY and STREET, *loc. cit.*
11. ASTBURY, W. T. and BELL, F. O., *Nature*, 147, 697 (1941)
12. LEMIN, D. R. and VICKERSTAFF, T., *Fibrous Proteins, op. cit.*, p. 129
13. RUDALL, K. M., *Fibrous Proteins, op. cit.*, p. 15
14. MERCER, E. H., REES, A. L. G. and FARRANT, J. L., *Nature*, 157, 589 (1946); 159, 535 (1947); *Austral. J. Exp. Biol. & Med. Sci.*, 24, 147, 175 (1946); *Anal. Chem.*, 20, 90 (1948)
15. HARRISON, W., *Am. Dyestuff Reptr.*, 27, 393 (1938)
16. ASTBURY, W. T., *Proc. Roy. Soc.*, 134B, 303 (1947)
17. SZENT-GYÖRGYI, A., *Acta Physiologica Scandinavica*, 9, Supp. No. XXV (1945); *Nature of Life. A Study of Muscle*, Academic Press, Inc., New York (1948)
18. SZENT-GYÖRGYI, A., *J. Coll. Sci.*, 1, 1 (1946)
19. SNELLMAN, O. and ERDÖS, T., *Nature*, 161, 526 (1948)
20. BERNAL, J. D. and CROWFOOT, D., *Nature*, 133, 794 (1934)
21. CROWFOOT, D., *Proc. Roy. Soc.*, 164A, 580 (1938)
22. SANGER, F., *Nature*, 162, 492 (1948)
23. SENTI, F. R. and WARNER, R. C., *J. Amer. Chem. Soc.*, 70, 3318 (1948)
24. BOYES-WATSON, J., DAVIDSON, E. and PERUTZ, M. F., *Proc. Roy. Soc.*, 191A, 83 (1947)
25. PERUTZ, M. F. and WEISZ, O., *Nature*, 160, 786 (1947)
26. LUNDGREN, H. P., *J. Am. Chem. Soc.*, 63, 2854 (1941)
27. LUNDGREN, H. P., *The Vortex* (Bulletin of the California Section of the American Chemical Society), IV, 231 (1943); WARD, W. H., HIGH, LORETTA M. and LUNDGREN, H. P., *J. Polymer Res.*, 1, 22 (1946); LUNDGREN, H. P., *Textile Res. J.*, 15, 335 (1945)
28. NUTTING, G. C., SENTI, F. R. and COPLEY, M. J., *Science*, 99, 328 (1944); *J. Phys. Chem.*, 49, 192 (1945)
29. ASTBURY, W. T., *Proc. Roy. Soc.*, 134B, 303 (1947); ASTBURY, W. T. and BELL, F. O., *Nature*, 145, 421 (1940)
30. BEAR, R. S., *J. Am. Chem. Soc.*, 66, 1297 (1944); 64, 727 (1942); KRATKY, O. and SEKOVA, A., *J. Makromol. Chem.*, 1, 113 (1943)
31. WRIGHT, B. A., *Nature*, 162, 23 (1948)
32. LLOYD, D. J. and ROBERTSON, M. E., *Nature*, 132, 974 (1933); 133, 102 (1934)
33. HIGHBERGER, J. H. and KERSTEN, H. J., *J. Am. Leather Chem. Assn.*, 33, 16 (1938)
34. HIGHBERGER and KERSTEN, *op. cit.*, p. 389

35. SPONSLER, O. L., BATH, J. D. and ELLIS, J. W., *J. Phys. Chem.*, **44**, 996 (1940)
36. PROCTER, H. R. and WILSON, J. R., *J. Chem. Soc.*, **109**, 307 (1916)
37. BERGMANN, M. and JACOBI, B., *Kolloid-Z.*, **49**, 46 (1929); KATZ, J. R., *Angew. Chem.*, **40**, 1443 (1927)
38. ASTBURY, W. T. and BELL, F. O., *Nature*, **141**, 747 (1938)
39. ASTBURY, W. T., *Nature*, **157**, 122, 124 (1946)
40. STANLEY, W. M., *Science*, **81**, 644 (1935); *J. Phys. Chem.*, **42**, 55 (1938)
41. WYCKOFF, R. W. G. and COREY, R. B., *J. Biol. Chem.*, **116**, 51 (1936)
42. LORING, H. S., LAUFFER, M. A. and STANLEY, W. M., *Nature*, **142**, 84 (1938); TAKA-HASHI, W. N. and RAWLINS, T. E., *Phytopath.*, **37**, 73 (1947)
43. BEST, R. J., *Nature*, **139**, 628 (1937)
44. STANLEY, W. M. and ANDERSON, T. F., *J. Biol. Chem.*, **139**, 325, 339 (1941)
45. FRAMPTON, V. L., *Science*, **95**, 232 (1942)
46. OSTER, G., DOTY, P. M. and ZIMM, B. H., *J. Am. Chem. Soc.*, **69**, 1193 (1947)
47. BERNAL, J. D. and FANKUCHEN, I., *J. Gen. Physiol.*, **25**, 147 (1941)
48. LANGMUIR, I., *Proc. Phys. Soc.* (London), **51**, 592 (1939)
49. MIRSKY, A. E. and PAULING, L., *Proc. Nat. Acad. Sci.*, **22**, 439 (1936)
50. ROTHEN, A. and LANDSTEINER, K., *Science*, **90**, 66 (1939)
51. DANIELLI, J. F., DANIELLI, M. and MARRACK, J. R., *Brit. J. Exp. Path.*, **19**, 393 (1938)
52. PAULING, L., *J. Am. Chem. Soc.*, **62**, 2643 (1940)
53. BOYD, W. C. and BEHNKE, J., *Science*, **100**, 13 (1944)

ARTIFICIAL POLYMERS. RESINS AND PLASTICS

Definitions

High polymers are derived from simple molecules or atoms called monomers, either by self-addition or by elimination of some small molecules such as water, or in a few cases by taking up such small molecules. Indeed it is the rule rather than the exception that some ordinary molecules such as an acid must be taken up to complete the molecular chain and that this end group has a significant effect upon the properties of the polymer. High polymers may be inorganic or organic, and a bridge between inorganic and organic polymers is provided by the organo silicone polymers.[1] Among inorganic polymers are sulfur, selenium, phosphorus, polyphosphonitrilic chloride and fluoride, metaphosphates and metaborates. Strictly speaking, some of the silica compounds should be included. Several authors would include soaps (cf. Chap. 19, page 240), but since their molecular weight varies readily and reversibly with mere change of concentration, like association colloids, this would seem to debase the term "polymer."

It is the practical importance of polymers as plastics that has brought them so conspicuously to the attention of the public in recent years. Clay is the ancient example of a thermosetting plastic, resulting in a permanently insoluble rigid state, but it is generally excluded from consideration under the topic. Glass is another familiar example of a thermoplastic which softens merely by the action of heat and may then be molded or formed to shape, and the process can be repeated indefinitely. But glass, too, is excluded from the modern definition of "plastic." Shellac in the form of sealing wax is familiar for its thermoplastic properties, but like bitumen, pitch, asphalt, and other natural

gum materials it too is not now counted a plastic. Likewise, the materials that are predominantly used in the lacquer and coating, film, and synthetic fiber industries are not classed as plastics. However, this enumeration serves to call attention to the enormous group of materials, most of which are high polymers and many of which are outstanding examples of colloid behavior. Their enormously developed surfaces lend themselves to all sorts of modifications; it is sufficient to recall the use of soap and surface active agents, dyes and other substances which dominate the technology of any one textile material.

"Modern plastics" is an apt designation because the commercial plastics are a development of only the last decades. True, Nobel in 1830 plasticized nitrocotton to form explosives, and J. W. Hyatt in 1869 plasticized cellulose nitrate with camphor to produce films and celluloid. Chardonnet in 1885 introduced "artificial silk," but even in 1912 some of the commercial products were almost as sensitive as candy to moisture; now rayons have been improved out of recognition. Spitteler in 1900 produced galalith from milk curd treated with formaldehyde; and in 1909 Leo Baekeland produced the first modern thermosetting resins from phenol and formaldehyde. Later urea formaldehyde was introduced, giving a colorless plastic which could be given any desired color. Then Talbot produced the thermosetting plastics from melamine and formaldehyde.

The distinctive feature of high polymers is a long chain held together by *primary valences* as in cellulose, the rubbers, and the proteins already discussed. The chain holds together under diverse conditions and in general is as difficult to break as to synthesize. This emphasizes the distinction from *association colloids*, such as soaps or ferric hydroxide, whose primary particles are made up, in some cases reversibly, from ordinary, small molecules by relatively weak forces, such as those of van der Waals and hydrogen bonds. However, these primary particles, like the high polymer molecules, may then further associate or aggregate to form larger structures such as ramifying aggregates or jellies.

It has recently been proposed to include aluminum soaps under the heading of high polymers [2] because they may string together by coordination of the aluminum atoms. However, they are primarily association colloids whose degree of association or "polymerization" depends upon the concentration of the solution and adjusts itself to that concentration [3] and other conditions. Thus they behave like sodium soaps in which such coordination is an unlikely mechanism

but in which other forms of association occur, such as that leading to jellies of the high sodium soaps in aromatic solvents.[4]

Staudinger distinguishes between typical association or micellar colloids, such as soap, and polymeric colloids which, however, can also associate and form micelles. If a polymer chain contains only 20–100 units, he terms it a hemicolloid; up to 1000 units, a mesocolloid; and above that a eucolloid. He indicates that the longest unbranched thread molecules are about 10,000 Å long, resulting in lyophilic colloids which swell with solvents and give viscous solutions at low concentrations.

Examples of High Polymers

The simplest high polymer is the polythene formed from ethylene by opening and joining the double bonds, an example of true polymerization which occurs in the presence of suitable catalysts under 12,000 atmospheres pressure:
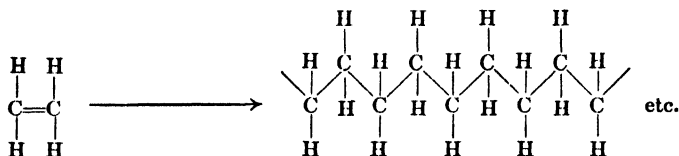
where the end valences have to pick up a suitable small molecule for termination, as already discussed. The result is a long chain of over 1000 units, which naturally fits other similar chains and therefore readily crystallizes into a sharp x-ray powder pattern, or, if stretched, an excellent fiber pattern. The fibers are orthorhombic, with unit cell dimensions a = 7.40, b = 4.93, and c = 2.52 Å, the latter in the direction of the fiber axis. Such fibers are of course very strong. As excessively long paraffin hydrocarbons they are excellent insulators, indifferent to water, tough, flexible, and of great impact strength.

If in ethylene one or more of the hydrogens of the vinyl group are substituted, a correspondingly substituted long chain is produced. Thus two CH_3 groups give polyisobutylene or Vistanex (a saturated hydrocarbon that cannot be vulcanized unless admixed with a little isoprene or butadiene); $COOCH_3$ gives acryloid; a CH_3 and $COOCH_3$ give methyl methacrylate or lucite; Ac gives polyvinyl acetate or vinylite A (which if then saponified yields polyvinyl alcohol, with all the OH groups on one side of the chain, hydrogen bonding them to the neighboring chain); Cl gives polyvinyl chloride or koroseal; C_6H_5 gives

polystyrene (partly branched and giving no x-ray pattern, with molecular weights up to 1 million);

$$HC \overset{H}{\underset{O}{\overset{|}{—}}} \overset{H}{\underset{}{\overset{|}{C}}}H \text{ gives } HO \left[\overset{H}{\underset{H}{\overset{|}{C}}} \overset{H}{\underset{H}{\overset{|}{C}}} —O \right]_m H$$
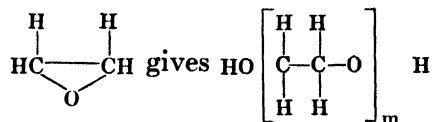
a polyglycol (the lower polymers being liquids and the higher ones waxes; if the terminal groups have polar compounds affixed they form detergents). The heat involved in polymerization is between 20 and 30 kilocalories per monomer.

Another series of the polymers is formed from butadiene and its chlorine and methyl derivatives; that is, from divinyl compounds. Other high polymers are co-polymers from the simultaneous polymerization of two kinds of monomers, as in the well known case of the synthetic rubbers Buna S or Buna N.

Then there is a long series of *condensation polymers* in which water or another small molecule is split off to leave the primary chemical bond. Examples are Bakelite or the glycerol and phthalic anhydride which gives glyptal. Both are cross linked three dimensionally, the latter due to the polyfunctional glycerol. Most of these reactions were worked out by organic chemists in the 19th century but were only recently turned to account.

The most famous of the recent reactions is that developed by W. H. Carothers [5] involving the condensation of diamines with dibasic acids yielding nylons. A check list of forty-nine of the more common polymers and their trade names is given by Powers.[6]

The Production of High Polymers [7]

There are three conditions under which polymerization is usually carried out:

1. The pure monomer is treated with a catalyst such as sodium or benzoyl peroxide.

2. The polymerization is carried out in solution giving a precipitate of the polymer.

3. The method affording the greatest control and of most interest from the standpoint of colloid science, now adopted where possible, is to carry out the polymerization in emulsions. For example, a mixture of three parts of butadiene and one of styrene is wholly or partially emulsified in a soap or detergent solution in the presence of a persulfate

or ferricyanide, and to control the tendency toward branching, modifiers such as mercaptans are added. The polymerization begins in the monomer solubilized in the lamellar micelles or in any interface lined with oriented soap molecules. The polymer, being of too high molecular weight, is not solubilized and is rejected, but is coated with detergent molecules. The later stages of the reaction proceed from such polymer particles.

With increasing molecular weight or degree of polymerization (D.P.) there is an increase in melting point, viscosity, insolubility, tensile strength (tending to a limit on reaching 1000 D.P.), impact strength, density (slight), and hardness.

At low temperatures a polymer tends to be brittle and may form a glass or form crystallites. As the temperature is raised, it tends to become plastic, ductile, and somewhat elastic. At still higher temperatures many thermoplastics become viscous liquids. The properties of the polymer are modified by plasticizers or high boiling solvents. Some polymers may be hardened by suitable additions, as for example casein by alum or formaldehyde. The thermosetting plastics are those which form cross-linkages on heating and therefore presumably set like fired clay.

The Properties of Plastics

The presence or absence of branching and, still more important, *cross-linking between chains* is determinative of the properties of the polymer. Houwink has given a complete review of the explanations of elasticity in terms of bondings.[8]

Short, linear polymers are usually soluble, but branched polymers less so. The effect of cross-linking is often extreme; for example, even only one part of divinyl benzene in 50,000 parts of styrene results in sufficient cross-linkages to make the polymer insoluble, although still extensible. With one part of divinyl benzene in a thousand of styrene, the polymer cannot even swell in solvents.

Cross-linking therefore makes polymers rigid (as in fully vulcanized rubber or ebonite), insoluble and resistant to liquids, infusible, dense, strong, resistant to abrasion, but often brittle. A resin is the result of cross-linking, as in the phenol formaldehyde resins. In the absence of cross-linking the polymers may be either linear and fibrous or rubbery. Side chains may interact to stiffen the structure, or they may prevent close packing or crystallization through steric hindrance, or, if highly indifferent, they may give a rubbery, elastic product.

Thus the properties of plastics are determined (1) by the chemical nature of the monomers, the most indifferent of which are the hydrocarbons; (2) the ease of packing or crystallization of the chains; side chains and branching reducing this possibility; (3) the average chain length; (4) the intrinsic flexibility of the chains themselves, which is reduced by acid groups, by double bonds, by ring formation, but most by (5) cross-linking; (6) by the magnitude and distribution of the intermolecular forces. In many cases, association of neighboring chains and side groups, formation of complexes, and dissociation of dissociable groups add considerably to the colloidal phenomena produced.[9] The softening temperature of a plastic is increased by cross-linking or by reducing the flexibility of the chain, whereas the side groups act as spacers or plasticizers and give rubber properties, especially if there are large amorphous regions in the plastic.

An exceptionally clear discussion of these relationships has recently been given by Mark.[10] The properties of plastics that have given them such universal importance are first that they are readily molded, and thus the most intricate designing may be produced and reproduced, as for example gramophone records or radio parts. This is an immense saving of labor. Next, they are of very low density, say half that of aluminum. Their cost is now comparable with an equal volume (not weight) of metal. Their strength and toughness, volume for volume, can be comparable with that of metals. They cover the whole range of color, from a transparency greater than that of glass, as in lucite or polystyrene, to black, opaque Bakelite. If properly chosen, they exhibit great permanency. For example, polystyrene is unaffected by air or water; rubber is attacked by gasoline, but neoprene is not. All of these general properties are additional to the specific properties of the various plastics.

The technology of plastics has been reflected in powder metallurgy where powdered metals are molded into coherent masses. Here, too, many colloid problems are involved.[11]

The Process of Formation of Polymers

The formation of a high polymer from a monomer is not the result of a single collision but of successive collisions, a growth reaction which depends upon pressure, temperature, solvent, acids, catalysts, and light. The first step is the opening of the double bond and the formation of the activated molecule or nuclei — a slow process. The activation energy is between 20,000 and 30,000 calories per molecule of mono-

mer, which indicates the dependence upon temperature. Next is the rate of growth, which is thousands of times quicker. Lastly comes the termination of the growth of the chain, usually by collision with another monomer, with a solvent molecule, with an impurity, or with the wall. Such a collision can activate, inactivate and stop, branch the chain or start another chain, or join the ends of a chain together to form a ring. These processes have an activation energy of only 4000 to 8000 calories and are therefore less dependent upon temperature and fall off with increasing size of the polymer. Sometimes the termination occurs through formation of a double bond by isomerization.

In contrast, polycondensation is a relatively slow, steady stepwise reaction with elimination of some small molecule such as water, ammonia, or hydrochloric acid at each linkage. Such a condensation is not reversed by heat as is a thermoplastic polymer but requires chemical reactions which may be difficult or incomplete. An example is the breakdown of cellulose to form sugar.

X-Ray Structure

The x-ray structure of high polymers and plastics may be illustrated by Figures 25·1 and 25·2, taken from the results of Bunn.[12] Figure 25·1 shows how perfectly the simple straight chains of polyisobutene may pack and crystallize, since they are unbranched and devoid of side chains. In contrast, Figure 25·2 shows the imperfect crystallization of drawn nylon fibers made from hexamethylenediamine and adipic acid. Before drawing, the nylon is disordered, plastic, and readily extended and has little tensile strength. After being stretched it is flexible, elastic, and has high tensile strength.

In the drawing of nylon, the exterior of the fiber is oriented more quickly and more perfectly than the interior, as is shown in Figure 25·3, which gives a series of micro x-ray diagrams taken by Fankuchen and Mark first through the center of the filament and secondly along the outside surface.

Fibers having micellar structure have small luster, swell a great deal, can be easily dyed, but are not fast to rubbing, and usually have low tearing and wet strength but high total stretch and good resistance to creasing. Fibers having pure chain structures have glassy or metallic luster, do not swell much, are difficult to dye, and have high tearing and wet strength but low total stretch and poor resistance to creasing. Most fibers are intermediate between these two extremes.
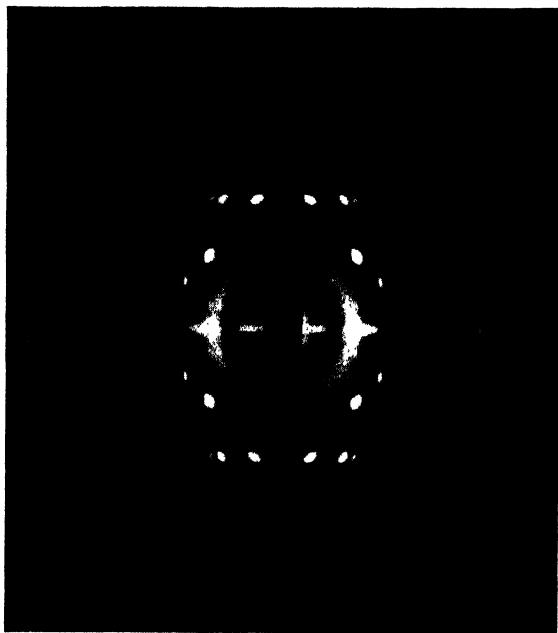
FIGURE 25·1 X-ray diffraction photograph of the straight chain, polyisobutene. Taken from Bunn, *J. Chem. Soc.*, 297 (1947), by courtesy of the author.
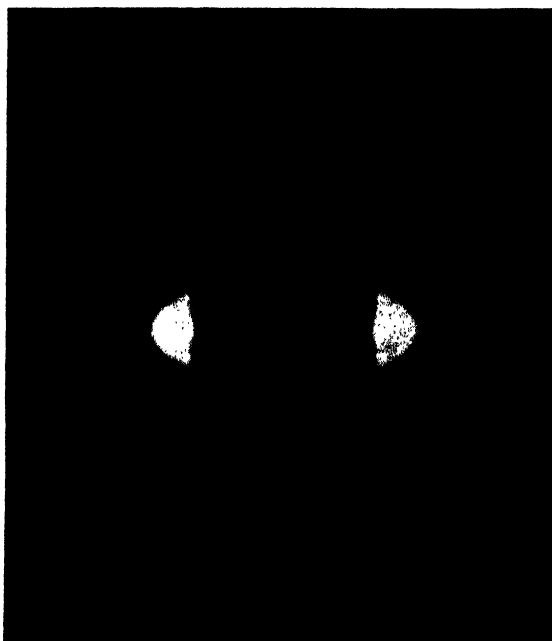
FIGURE 25·2 X-ray diffraction photograph of drawn nylon fibers (polyhexamethyleneadipamide), showing imperfect crystallization. Taken from Bunn, *J. Chem. Soc.*, 297 (1947), by courtesy of the author.

Spreading or Films

The spreading of high polymers upon water has been investigated by Crisp.[13] Here again, as with proteins, the spreading depends upon the hydrophilic side groups which are attracted to the water, the hydropho-

FIGURE 25·3 Series of micro x-ray diagrams, indicating the mechanism of drawing of nylon. Taken from Fankuchen and Mark, *J. Appl. Phys.*, 15, 364 (1944), by courtesy of the authors.
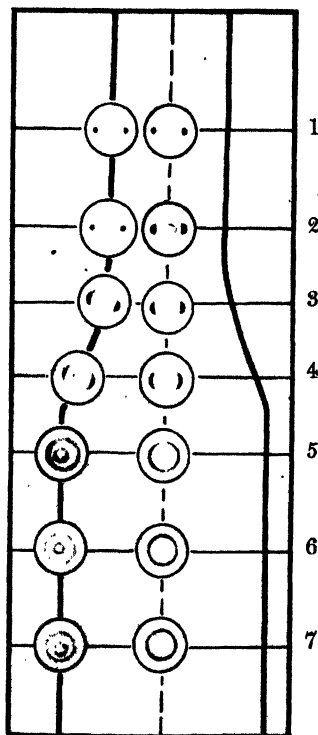
bic groups being repelled. The spreading is hindered by cross-linking, for example, or by high cohesion between the polymer molecules. Some linear polymers form fluid monomolecular films which can be compressed into thicker "overfilms" and then into thick layers with the properties of the bulk polymer. When the intermolecular forces of the polymer are high, it is only at the very lowest surface pressures that the films can persist without collapsing. "Short side chains result in strong intermolecular forces and coherent films; flexible side chains reduce these forces, giving a more expanded type, while large side chains introduce secondary cohesion." The films dissolve when a sufficient proportion of the groups therein are ionized. Polyamides form very coherent films due to bonding of the polypeptide link. The "overfilm" in many cases is three-dimensional.

REFERENCES

Resins and Plastics

ALFREY, TURNER, "Mechanical Behavior of High Polymers," Vol. VI of *High Polymers*, Interscience Publishers, Inc., New York (1948)

BARRON, H., *Modern Plastics*, John Wiley & Sons, Inc., New York (1945)

BARRY, T. H., *Natural and Varnish Resins*, D. Van Nostrand & Co., Inc., New York (1933)

BAWN, C. E., *The Chemistry of High Polymers*, Interscience Publishers, Inc., New York (1948)

BELL, L. M. T., *The Making and Moulding of Plastics*, Chemical Publishing Co., Brooklyn (1938)

BURK, ROBERT E., et al., *Polymerization and Its Application in the Fields of Rubber, Synthetic Resins and Petroleum*, A.C.S. Monograph # 75, Reinhold Publishing Corp., New York (1937)

CARSWELL, T. S., "Phenoplasts, Their Structure, Properties and Chemical Technology." Vol. VII of *High Polymers*, Interscience Publishers, Inc., New York (1947)

DUBOIS, J. H., *Plastics*, American Technical Society, Publishers, Chicago (1943)

ELLIS, C., *Chemistry of Synthetic Resins*, Reinhold Publishing Corp., New York (1940)

Faraday Society, "A general discussion on the phenomena of polymerisation and condensation," *Trans. Farad. Soc.*, 32, 1–227 (1936); particularly, CAROTHERS, W. H., "Polymers and Polyfunctionality," p. 39

FLECK, H. R., *Plastics, Scientific and Technological*, Chemical Publishing Co., Brooklyn (1945)

FUOSS, R. H., et al. (17 authors), *High Polymers*, Annals N. Y. Acad. Sci., pp. 263–443. Published by the Academy, New York (1943)

MARK, H. and PROSKAUER, E. S., *The Science of Plastics: A Comprehensive Source Book Based on Original Literature for 1942–46*, Interscience Publishers, Inc., New York and London (1948)

MARK, H. and RAFF, R., "High Polymeric Reactions, Their Theory and Practice," Vol. III of *High Polymers*, Interscience Publishers, Inc., New York (1941)

MARTIN, S. R. W., *Synthetic Resin Chemistry for Students*, Chapman & Hall, Ltd., London (1947)

PLASTES, *Plastics in Industry*, Chemical Publishing Co., Brooklyn (1941)

POWERS, PAUL O., *Synthetic Resins and Rubbers*, John Wiley & Sons, Inc., New York; Chapman & Hall, Ltd., London (1943)

ROBINSON, H. A., *High Polymer Physics*, Chemical Publishing Company, Inc., New York (1948)

ROWELL, H. W., *The Technology of Plastics*, Plastics Press, Ltd., Ludgate Hill, London (1936)

SCHEIBER and SANDIG, *Die kunstliche Harze*, Weisz Verlag, Stuttgart (1943)

SCHMIDT, A. X. and MARLIES, C. A., *Principles of High-Polymer Theory and Practice*, McGraw-Hill Book Co., New York (1948)

Society of Plastics Industry Handbook, Society of Plastics Industry, New York (1948)

TAMMANN, G., *Der Glaszustand*, Verlag Leopold Voss, Leipzig (1933)

WAKEMAN, R. L., *The Chemistry of Commercial Plastics*, Reinhold Publishing Corp., New York (1947)

WEIL, B. H. and ANHORN, VICTOR J., *Plastic Horizons*, The Jaques Cattell Press, Lancaster, Pa. (1944)

WINDING, C. C. and HASCHE, R. L., *Plastics Theory and Practice; The Technology of High Polymers*, McGraw-Hill Book Co., Inc., New York and London (1947)

Text References

1. ROCHOW, E. G., *Introduction to the Chemistry of the Silicones*, John Wiley & Sons, New York (1946); *Chem. Eng. News*, **23**, 612 (1945); COLLINGS, W. R., *Chem. Eng. News*, **23**, 1616 (1945)

2. GRAY, V. R., *Trans. Farad. Soc.*, **42B**, 197 (1946); MCGEE, C. G., *J. Am. Chem. Soc.*, **70**, 1053

3. MCBAIN, J. W. and WORKING, E. B., *J. Phys. Coll. Chem.*, **51**, 974 (1947)

4. MCBAIN, J. W. and MCCLATCHIE, W. L., *J. Phys. Chem.*, **36**, 2567 (1932)

5. His classical researches are collected in Vol. 1 of *High Polymers*, Interscience Publishers, Inc., New York (1939)

6. POWERS, P. O., *Chem. Eng. News*, **24**, 2784 (1946)

7. FLORY, P. J., *Chem. Rev.*, **39**, 137–197 (1946)

8. HOUWINK, R., *Kolloid-Z.*, **77**, 183 (1936)

9. WALTER, G., *Trans. Farad. Soc.*, **32**, 377 and 402 (1936)

10. MARK, H., *Chem. Eng. News*, **27**, 138–142 (1949)

11. RHINES, F. N., *Powder Metallurgy Bull.*, **3**, 28 (1948)

12. BUNN, C. W., *Proc. Roy. Soc.*, **191A**, 297 (1947)

13. CRISP, D. J., *J. Coll. Sci.*, **1**, 49, 161 (1946)

CLAY

About four per cent of the igneous rocks of the earth's crust have been weathered in the course of geological ages to clays, shales, and surface soils. These are mixtures of finely divided compounds. As a result of this immensely prolonged exposure and weathering they would be expected to be completely inert. However, their behavior is colloidal because of the fineness of the primary particles, the dominance of their surface reactions, and the prominence of their tendency to associate and to disaggregate. These intimate mixtures of fine materials contrast with the single chemicals which in general formed the basis of colloidal particles discussed in the previous chapters.

Many volumes would be required to treat adequately of clays and their colloidal behavior. A descriptive report of their occurrence and properties, as known up to 1927, was published by the geologist Ries [1] which is still of value although it just antedates knowledge of the structure of clays which resulted from application of x-ray techniques. Other books deal with the structure and particle size of soil.[2]

Definition and Particle Size

From the standpoint of soil science all inorganic particles of sufficiently small size are called "clay." Atterberg in 1905 placed 2 microns as the upper limit of clay particle size, choosing this as the limit at which Brownian movement becomes inappreciable. This limit was adopted in 1913 by the International Society of Soil Standards. Anything less than 5 microns was the standard adopted by the U.S. Bureau of Soils from 1896 to 1937, and by the A.S.T.M. Test Procedure D–422

in 1939. The A.S.T.M. definition of "colloidal clay" is that which is less than 1 micron. This is usual in soil science.

These dimensions are not actual dimensions but are derived from observing the rate of settling of clay particles in water or aqueous solutions and calculating the size of corresponding hypothetical solid spherical particles of the same density that according to Stokes' law would have settled at that rate. The formula used is:

$$d = \frac{30\eta L}{980 \ (w - w_1) \ t}$$

where

d = spherical diameter of particles in mm,
η = the viscosity of the suspension in poises,
w = the density of the solid particle,
w_1 = the density of the fluid,
L = the depth of fall in cm,
t = time of fall in minutes.

Hence, a thin plate might be many times larger in diameter than the hypothetical sphere.

There are many dozens of procedures for carrying out such sedimentation analysis. A history of these, going back to Hippocrates in 400 B.C., has been given by Krumbein,[3] but the earliest true mechanical analysis dates from 1804. Often the sample is boiled before testing (Sir Humphry Davy, 1815). Usually some peptizing agent is employed to separate the clay particles from each other. The most commonly used are ammonia, hydrogen peroxide, sodium carbonate, sodium hydroxide, sodium oxalate, sodium citrate, sodium metaphosphate, sodium silicate, and gum arabic. The writer's experience is that with some clays wholly different results are obtained by these various procedures and peptizers. With some clays fine grinding does not increase the percentage of fine particles so obtained. Indeed with some, even passing the impure clay through a colloid mill produces no larger proportion of particles less than 1 micron. According to Wiley, a fine, natural clay might contain about 75% of "clay"; heavy clay soils, 45%; ordinary loamy soils, 15%.

A whole literature has grown up around the subject of sedimentation. Generally the procedure or peptizer giving the largest proportion of fine particles is taken as most significant. The dominant phenomena here are aggregation and disaggregation. An example of this colloidal

behavior is that the rivers of the earth carry the colloidal parts of top soils and clays more readily if they are slightly alkaline. Where they enter the sea, the particles are coagulated by the calcium and magnesium salts and produce the great deltas.

Discrimination between particles much smaller than 1 micron can only be made with the use of a centrifuge, used from time to time since 1895, but now superseded by the Sharples supercentrifuge.[4] This has become important since it has been found that fractionation into sizes smaller than 0.1 micron frequently separates quite different mineral species.[5] According to soil science *sand* is the name given to particles more than 20 microns in diameter and *silt* to particles lying between this and clay, but it has been shown that finer particles of silica may extend down as low as 0.1 micron.

One-micron spheres of the same density, 2.6, as clay would sediment about 8 cm in 24 hours. Hence all particles suspended above this depth are of smaller equivalent diameter. The finest particles of clay remain suspended in distilled water for months or years. Rembest ball clay, an English kaolin, remains almost quantitatively in suspension even after six months if the water is made slightly alkaline with sodium hydroxide.

The phenomena of sedimentation are not always completely simple. For example, when different weights of yellow bentonite clay, such as 10.5 gm, 5.3 gm, 1 gm, and 0.1 gm, respectively, were each placed in 153 gm of distilled water, only the 1-gm sample stayed suspended; the others all settled out in a number of hours. At pH 7.0 much more suspension was exhibited in all cases except with the sample weighing 5.3 gm.

The foregoing raises again the definition of colloidal condition, which is so often defined by size limits. One micron is often taken as the lower practical limit of the resolving power of a fairly good microscope, and anything smaller is termed "colloidal." Ten Ångström units was arbitrarily selected by Ostwald as an approximate lower limit in order to exclude ordinary small chemical molecules. However, it has been made abundantly clear that there is no sharp boundary at either end in connection with any intrinsic property. This is particularly evident in the case of highly elongated particles or very thin disc-like particles. That is why Bancroft includes in the colloid domain "bubbles, drops, grains, filaments, and films, because in each of these cases at least one dimension of the phase is very small," [6] and hence surface properties dominate bulk properties.

Classification of Clays among Colloids

Attempts were made at the beginning of this century to put colloids into two groups according to coagulation or flocculation by electrolytes or heat. However, the difficulty arises that many colloids sensitive to electrolytes are insensitive to heat and vice versa. Hardy in 1899 grouped together materials with irreversible and those with reversible coagulation. Unfortunately such coagulation depends upon heat, upon the nature of the electrolyte, and even for a given colloid upon concentration. Müller in 1903 separated those resembling suspensions, such as gold sol or arsenic trisulfide, from those resembling solutions, such as albumin and many high polymers. Bechhold in 1904 put inorganic colloids into his second order, keeping such substances as albumin in his first order. Perrin in 1905 termed his groups hydrophobe, or easily coagulated, and hydrophile, or reversibly swelling. Noyes in 1905 gave the more elaborate description, "non-viscous, non-gelatinizing, but readily coagulable by salts, as distinguished from viscous, gelatinizing, colloidal mixtures not easily coagulated by salts." Zsigmondy in 1906 distinguished irresoluble and resoluble colloids.

Ostwald in 1907 attempted to put the matter on a theoretical basis, and for a time this met with great acceptance. He distinguished what he called suspensoids, whose particles were solids, from emulsions or emulsoids, whose particles were liquid. He identified the latter with such substances as albumin, soap, high polymers, and the like. Unfortunately, many finely divided emulsions and colloidal liquids such as mercury have the behavior which he regarded as typical of suspensoids, whereas ultrafiltration and many other data indicate that in general the colloidal particles of the so-called emulsoids are not liquid.

Although none of these proposed classifications can be consistently applied and although they by no means coincide, yet there is in general a distinction to be drawn between stable and unstable colloids, even if it is only a difference in degree. Frequently, however, a so-called unstable colloid is one with only a small amount of perhaps adventitious stabilizing agent, yet often such systems persist indefinitely if kept undisturbed and free from contamination.

Now where does clay come in these classifications? In general, it might be placed under the first mentioned group of the various authors. Ostwald included it specifically as a suspensoid. However, the coagulation can be reversed by washing out with noncoagulating or peptizing electrolytes, and all clays swell and many readily suspend in water.

TABLE 22

CLAYS AND ALLIED SUBSTANCES

Hydrated Silica	Brucite	Gibbsite
4 OH	6 OH	6 OH
4 Si	6 Mg	4 Al
6 O	6 OH	6 OH
	Thickness 4.73 Å	Thickness 4.86 Å

Halloysite
6 OH
4 Al
6 OH
4 OH
4 Si
6 O
Thickness 10.25 Å

Chrysotile Asbestos $Mg_3Si_2O_5(OH)_4$

Kaolinite $Al_2Si_2O_5(OH)_4$

China Clay = Kaolinite $Al_2O_3 \cdot 2\,SiO_2 \cdot 2\,H_2O$

6 OH
4 Al   47% SiO_2
4 O + 2 OH   40% Al_2O_3
4 Si
6 O
Thickness 7.2 Å

Idealized Montmorillonite
6 O
4 Si
4 O + 2 OH
4 Al
4 O + 2 OH
4 Si
6 O
Thickness 9.6 + Å
Water layer or layers of thickness 9.2–21.4

Hydrous Micas and Illite*
$Al_2O_3 \cdot 2\,SiO_2 \cdot \frac{1}{3}\,K_2O \cdot \frac{1}{3}\,H_2O$
Mica (Muscovite)
K
6 O
3 Si + Al   45% SiO_2
4 O + 2 OH   38% Al_2O_3
4 Al   12% K_2O
4 O + 2 OH   4.5% H_2O
3 Si + Al
6 O
K
Thickness 10.0 Å

$Al_2O_3 \cdot 4\,SiO_2 \cdot H_2O$
Pyrophyllite

tetrahedral	6 O	
	4 Si	67% SiO_2
	4 O + 2 OH	28% Al_2O_3
octahedral	4 Al	5% H_2O
	4 O + 2 OH	
tetrahedral	4 Si	
	6 O	

Thickness 9.6 Å
$(Al_2)(Si_4)O_{10}(OH)_2$
$(Mg_3)(Si_4)O_{10}(OH)_2$

Areal montmorillonite $(Al_{1.67}Mg_{.33})(Si_4)O_{10}(OH)_2 \cdot Na_{0.33}$
Hectorite $(Mg_{2.67}Li_{0.3})(Si_4)O_{10} \cdot (OH)_2 \cdot Na^+_{0.3}$
Mg, Beidellite, Saponite $Mg_3(Al_{0.3}Si_{3.7})O_{10} \cdot (OH)_2 \cdot Na_{0.3}Aq$
Beidellite $Al_2O_3 \cdot 3SiO_2 \cdot 4\,H_2O$ { 51% SiO_2 29% Al_2O_3 }
Nontromite $(Fe_2)(Al_{0.3}Si_{3.7})O_{10}(OH)_2$

Bentonite and Fuller's Earth are mostly Montmorillonites and Beidellite (and perhaps some hydrous mica interleaved), swelling if exchangeable base is Na+; nonswelling if Ca++ or Mg++ or H+. Florida Fuller's Earth is Attapulgite. Anauxite, $Al_2O_3 \cdot 3\,SiO_2 \cdot 2\,H_2O$, may be interleaved kaolinite and silica layers.

*Illite = Bravaisite = Ordovician Bentonite = Hydrous Mica = Potash-bearing Clay
2 K_2O · 3(MgO and FeO) · 8(Al_2O_3 and Fe_2O_3) · 24 SiO_2 · 12 H_2O
6%   3%   1.5%   25%   6%   51%   7%

The Mineral Species of Clay[7]

The mineralogy of the finely divided clay particles was very confused until about 1930 when it was recognized that nearly all are highly crystalline platelets, or occasionally fibers, belonging to a number of types of mineral species. Previously they had been considered, with few exceptions, to be amorphous, hydrous aluminum silicates or, occasionally, hydrous magnesium silicates, but generally of varying composition due to isomorphism and base exchange. Some natural deposits had been found sufficiently segregated to be recognized as mineral species. Such were kaolin, montmorillonite, and Wyoming-type bentonite.

The whole subject was placed upon a new foundation when it was recognized from x-ray analysis that there are two main types of structure of which halloysite and kaolinite, on the one hand, and montmorillonite and hydrous mica on the other hand, were the chief exemplars. In the kaolin type there are alternate layers of alumina and of hydrated silica. In the other type there are three layers like a sandwich in which the center layer is largely alumina and the outside layers are hydrated silica. The relations between these mineral types and their simple antecedents are given in Table 22. Such kaolin-type minerals as dickite, nacrite, and erdellite are omitted; so is the fibrous mineral attapulgite. Halloysite particles are acicular.

The structural arrangement of the 2-layer and of the 3-layer clays is better illustrated by Figure 26·1, taken from Hauser.[8]

A scale drawing by Jackson [9] of a montmorillonite-type clay is given in Figure 26·2, with a water layer 15 Å thick containing an ionic atmosphere. For the alkali ions two circles are drawn, the inner one being the anhydrous atom, the outer one representing a probable size for the hydrated ion as it exists in solution.

The first thing one notices about Table 22 and Figures 26·1 and 26·2 is the layer or platelike structure of these minerals. Mica is outstanding for its ready cleavage, so it is not surprising that water is often able to separate the micaceous clay particles by penetration into the spaces separating the two or three layers of bonded atoms. This is far more pronounced with the 3-layer clays, both of whose external surfaces are composed of silica, than with the 2-layer clays where silica and alumina alternate and are more firmly held together. In significant contrast with the highly developed cleavage of potassium mica, muscovite, the corresponding mineral, clintonite, may be mentioned where the divalent cation calcium replaces the monovalent potassium and
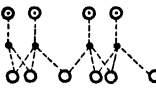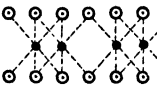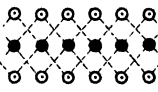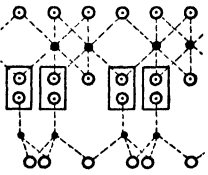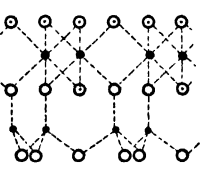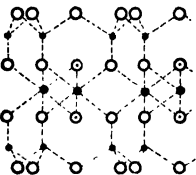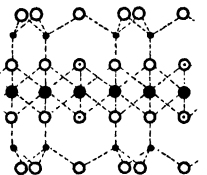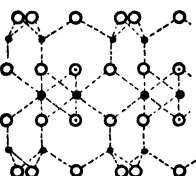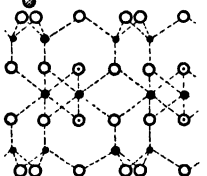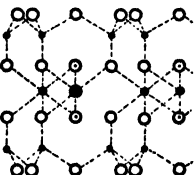
1 SILICON TETRAHEDRON		2 ALUMINUM OCTAHEDRON		3 MAGNESIUM OCTAHEDRON	
1O	−2	3OH	−3	3OH	−3
1Si	+4	1Al	+3	1Mg	+2
3O	−6	3OH	−3	3OH	−3

4 HYDRATED SILICA		5 GIBBSITE		6 BRUCITE	
4OH	−4	6OH	−6	6OH	−6
4Si	+16	4Al	+12	6Mg	+12
6O	−12	6OH	−6	6OH	−6

7 HALLOYSITE		8 KAOLINITE	
6OH	−6	6OH	−6
4Al	+12	4Al	+12
6OH	−6	4O + 2OH	−10
4OH	−4		
4Si	+16	4Si	+16
6O	−12	6O	−12

9 PYROPHYLLITE AND MONTMORILLONITE (IDEAL CASE)		10 TALC	
6O	−12	6O	−12
4Si	+16	4Si	+16
4O + 2OH	−10	4O + 2OH	−10
4Al	+12	6Mg	+12
4O + 2OH	−10	4O + 2OH	−10
4Si	+16	4Si	+16
6O	−12	6O	−12

11 NONTRONITE		12 MICA (ILLITE)	
6O	−12	1K	+1
4Si	+16	6O	−12
4O + 2OH	−10	3Si + 1Al	+15
4Fe^{+++}	+12	4O + 2OH	−10
4O + 2OH	−10	4Al	+12
4Si	+16	4O + 2OH	−10
6O	−12	3Si + 1Al	+15
		6O	−12
		1K	+1

13 MONTMORILLONITE (SUBSTITUTED)		KEY	
6O	−12	● − Si	
4Si	+16	● − Al, Fe^{+++}	
4O + 2OH	−10	● − Mg	
3Al + 1Mg	+11	○ − O	
4O + 2OH	−10	◉ − OH	
4Si	+16	◉ − K	
6O	−12		

FIGURE 26·1 Structural data of clay minerals, according to E. A. Hauser, *J. Am. Ceramic Soc.*, **24**, 183 (1941). The schematic drawings in the first columns represent the composition of the unit cell of the crystal lattice. All atoms are projected into one plane. The second column gives the number and type of atom or group in every lattice plane. The third column reveals the amount of available or needed electrons in every sheet. If the (+) and (−) in this column compensate each other the crystal is electrically neutral. If they do not, the lattice carries either a positive or negative charge. As the plus and minus charges are listed for every layer, it is easy to determine where the unbalance actually occurs in the lattice.
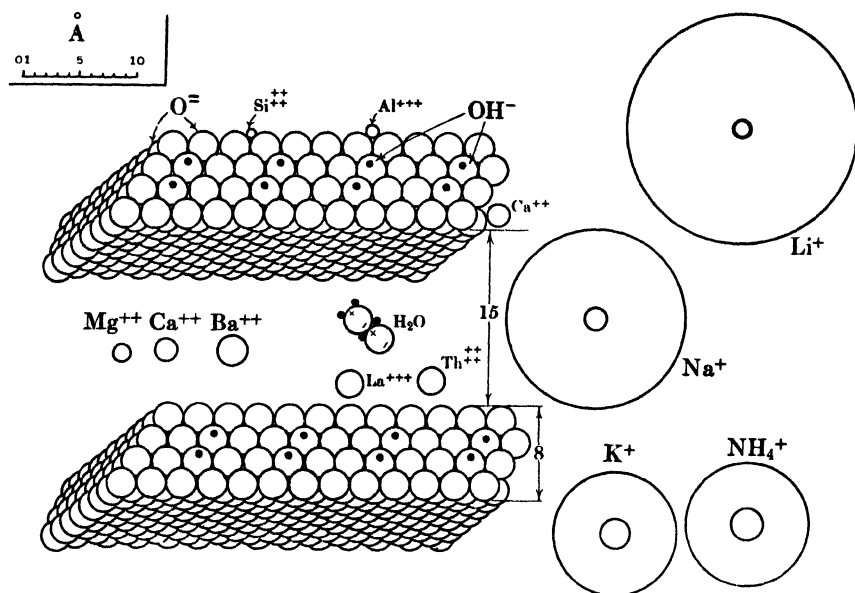
FIGURE 26·2 Scale drawing of a montmorillonite-type clay, according to M. L. Jackson, Ph.D. Thesis, Univ. of Wisconsin, 1939.

firmly binds the layers together so that cleavage is absent and water also cannot penetrate.

It is doubtful whether all the common clays can be included within the types recognized within Table 22, page 381. The situation is much complicated by isomorphous replacements, particularly in the 3-layer types, by not infrequent interlayering of different types in the same microscopic crystal, and by the variability of minor constituents due to base exchange. Type kaolinite is exceptional in that some deposits of china clay consist largely of this one material first recognized in 1780. The chemical analyses of kaolin quoted by Dana vary by several per cent in each of their components from the type formula, and some include several per cent of other oxides.

Wyoming-type bentonite differs significantly in composition and properties in different portions of the same deposit. Bentonite is a naturally occurring rock resulting from the weathering of volcanic ash. Its chief mineral species, montmorillonite, was characterized in 1847 solely on the basis of a chemical analysis of a French deposit which was not a bentonite and, judging by a more recent specimen of de Lapparent, was not pure but, according to its x-ray diagram, shows the presence of illite and particularly kaolin. Kaolin is a product of the weather-

ing of feldspar. The type beidellite, named in 1925, has been queried in recent years on the basis of thermal analysis. The hydrous micas now often include compositions that depart widely from the original type illite described by Grim in 1937. Finally, as will be seen, the properties of a clay are profoundly modified by base exchange, which involves chiefly minor constituents whose determination in the present state of the commercial art is difficult and not precise. It is not surprising then that even a careful chemical analysis, apart from extreme cases, gives but little information as to the behavior of a clay. Of course, it can settle such a point as whether there is enough iron to spoil the white color of ceramic ware after firing.

There are two kinds of isomorphous substitution recognized as occurring in the 3-layer lattice clay minerals. The aluminum of the middle layer can be progressively substituted by magnesium, lithium, manganese, or iron. These are the montmorillonites. If in addition, alumina substitutes for some silica in the outermost layer, a beidellite results. The hydrous micas, commonly occurring clays which resemble muscovite but contain far less potash, are sufficiently indicated in the table.

X-Ray Examination of Clays

X-ray examination of clays is almost entirely confined to powder photographs, or the far more rapid and convenient examination with the Norelco Geiger Müller X-Ray Spectrometer. Reference to Figures 26·1 and 26·2 shows that a change in the water layer causes a change in the long spacing normal to these layers, whereas the bonded layers themselves are fixed by their chemical composition. The thickness of the water layers is a function of relative humidity, of previous treatment, and of the solutions to which the clay has previously been exposed. Deep shales (compacted clays) have had most of the planar water pressed out by the weight of the overburden and frequently cause much trouble in tunneling and in drilling operations because they swell by absorption of water from drilling mud or from the surface when exposed. Such clays are called "heaving shales." Calcium or other polyvalent ions minimize this effect and are useful in preventing slides such as are encountered in the Panama Canal.

It is desirable to get clays into a standard state before x-raying them. In accordance with a suggestion of MacEwan,[10] they may be first treated with aqueous glycerin and then dried. Various organic liquids have been used with corresponding results.[11]

With glycerin, montmorillonite gives a long or basal spacing between 17.8 and 18 Å, as compared with the characteristic spacing of 7.1 Å (and a still stronger second order line of 3.58 Å) for the kaolinite clays and a diffuse spacing in the neighborhood of 10 Å for the hydrous micas (sharp for mica itself). Beidellite, illite, and pyrophyllite, the latter of which is nonexpanding, do not give the 18 Å spacing. However, it is strongest in hectorite. An estimation of the amount of montmorillonite in a given sample may be obtained by comparison of the peak at 18 Å obtained with the spectrometer, with that of another clay to which varying amounts of montmorillonitic clay have been added. However, the result may be affected by different degrees of orientation, size of crystallites, variable basal spacings in the same crystallite, or interlayering of different mineral species.

The three typical x-ray spacings mentioned above are shown in Figure 26·3, which includes a Louisiana illitic clay, a china or kaolinitic clay, and a Wyoming bentonite. The shorter spacings of these three clays are shown in Figure 26·4. Unfortunately, a number of strong lines are common to all clays, such as that at 4.46, that at 1.50, which is shared by calcite and near to one of quartz (and feldspar), and that at 2.56, shared by feldspar. Another line shared by clay and feldspar is that at 1.27 Å. A number of other minerals as well as halloysite give a spacing between 7.2 and 7.5 Å.

BENTONITE   17.8 Å

KAOLIN   7.1 Å

ILLITIC   10 Å
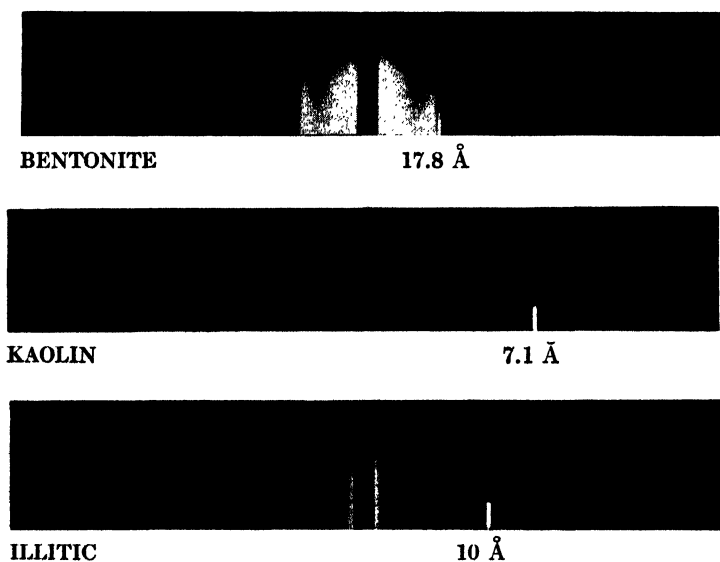
FIGURE 26·3 X-ray "long" spacings of a bentonite, a kaolin, and an illitic type of clay, showing typical, characteristic spacings.

BENTONITE

KAOLIN

ILLITIC
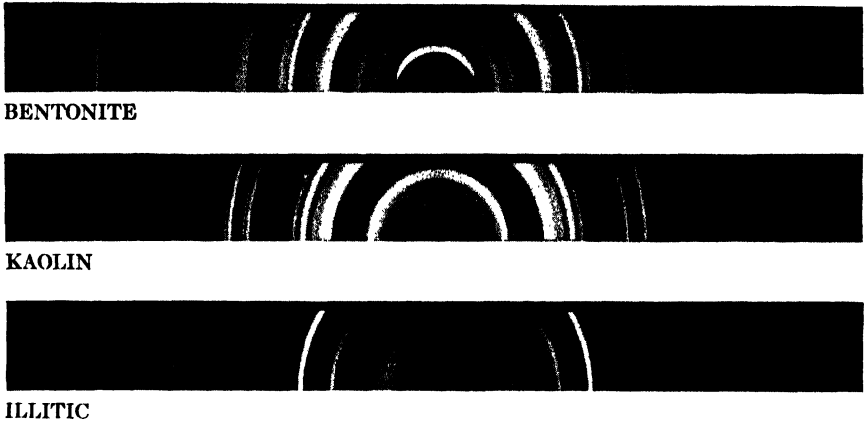
FIGURE 26·4 X-ray "short" spacings of the same three clays shown in Figure 26·3.

Petrographic examination is important for picking out admixed minerals such as zircon, quartz, and calcite and for noting a shard structure in special samples of bentonite. However, the discrimination of the finely divided particles of most clays by means of refractive index, etc., is difficult and sometimes impossible. For these reasons several supplementary methods of examination are required.

Thermal Analysis

Differential thermal analysis has become a most important method of investigating clays.[12] In this, the rate of heating of a sample of clay as compared with a sample of already heated clay or indifferent material is recorded against the temperatures at which such differences occur. The curves of three such thermal analyses are given in Figure 26·5, viz., the same illite as that in Figures 26·3 and 26·4, a kaolin, and a Wyoming bentonite. The illite contains not more than 5% montmorillonite, although the endothermic reaction at 330° C may indicate the presence of about 10% gibbsite.

The first dip in the curves, corresponding to absorption of heat, occurs when water is given off. Bentonite, of which the chief mineral species is montmorillonite, has a very large dip in the neighborhood of 150° C, corresponding to the large amount of planar water that is there removed, whereas kaolin has none. Illite and halloysite are intermediate. Montmorillonite decomposes with absorption of heat above 700° C followed by a smaller absorption of heat at about 850° C and a rise in the curve at about 900° C. Kaolins exhibit a deep dip at about

600° C, followed by a high sharp peak between 950 and 1000° C. Illite is least conclusively identified. A clay not containing large amounts of iron or magnesium and showing a dip between 500 and 600° C is recognized as illite unless a sharp peak near 1000° C shows it to be kao-
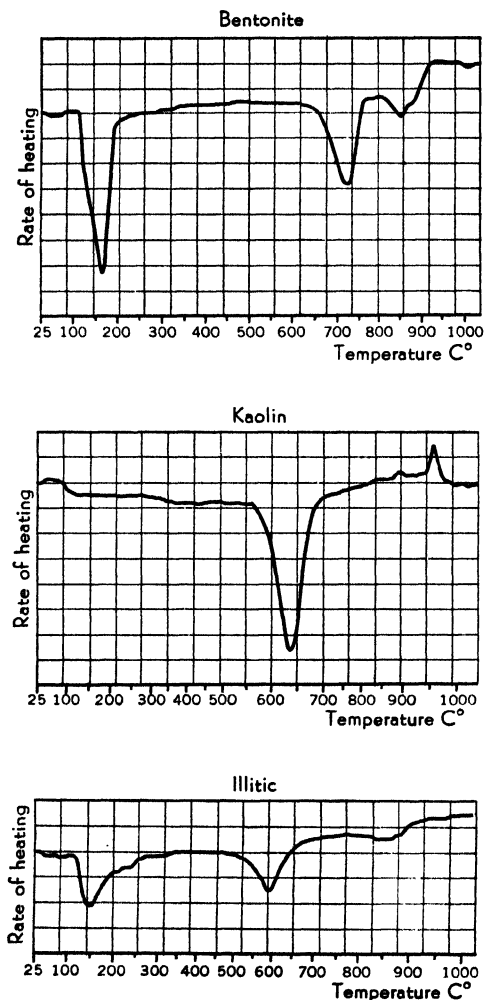
FIGURE 26-5 The curves resulting from thermal analyses of three typical clays.

linitic or admixed with kaolin. From the standpoint of thermal analysis an illite is almost any clay material that is not montmorillonite and not kaolin. Again, in this connection, Grim repeats the warning that chemical analysis, x rays, thermal analysis, if taken separately, are not a safe identification of clays.

FIGURE 26·6 Electron micrograph of kaolinite, unheated. Taken from Nelson, R. A., and Hendricks, S. B., *Soil Sci.*, 56, 291 (1943).

FIGURE 26·7 Electron micrograph of a Wyoming bentonite, 1-micron fraction. Taken from Alexander, J., *Colloid Chemistry*, by courtesy of Reinhold Publishing Corp.

Electron Microscope [13]

The electron microscope resolves the small particles of clay, revealing their shapes, and, by degree of transparency to electrons, indicates how far they have been broken down toward the ultimate two or three layers. Kaolin crystals tend to show rather sharp edges, and the crystals tend to remain at least partly intact, although they can be broken up. This is illustrated in Figure 26·6. On the other hand, the bentonite particles have very irregular edges and tend to separate into very thin platelets down to the ultimate three layers, as is shown by some of the mosaic of particles in Figure 26·7. Most of the particles lie between 0.1 and 1 micron, although some are smaller. Most of the particles of montmorillonites lie between 0.1μ and 200 Å.

Base Exchange

Base exchange results from the difference in valence of the trivalent aluminum substituting for tetravalent silica in the outer layer, or of divalent magnesium or ferrous iron substituting for trivalent aluminum in the inner layer. Each such substitution leaves one negative charge upon the lattice which has to be compensated by free positive ions upon or near its surface. If the external ion of montmorillonite is sodium, the result is a highly swelling clay of the Wyoming bentonite type. If it is hydrogen or calcium or magnesium or other polyvalent ion, the clay coagulates and swells very little. Even kaolinitic minerals show some base exchange at the edges and corners of the particles where the bonds are broken and left unsaturated,[14] but the amount of ion exchange in kaolin is very small as compared to that of the substituted 3-layer clays. A bentonite sorbs about ten times as much dye as a kaolinitic clay.

Thiessen has found that colloidal gold particles are adsorbed only on the edges of the platelike kaolin particles, but with micaceous particles they are adsorbed on the surface also. After kaolin is treated with various ions or cation active detergents, gold particles are adsorbed on its surface too.[15]

The number of milliequivalents of cations adsorbed per 100 gm of clay is known as the *base exchange capacity* or B.E.C. For montmorillonite clays the B.E.C. usually varies between 60 and 110 but may reach 150; for kaolinitic clays it lies between 3 and 15. Pyrophyllite is a nonswelling mineral that does not exhibit definite base exchange.

TABLE 23

EXCHANGEABLE BASES IN MILLIEQUIVALENTS PER 100 GRAMS OF CLAY

	Volclay	Panther Creek	Ohio Plastic	Kentucky Ball
Ca	22.0	64.7	5.2	7.1
Na	85.5	0.4	0.4	0.5
K	5.0	2.8	0.5	0.4
Mg	1.0	1.0	1.3	3.8
Actual Total	113.5	68.9	7.4	11.8
B.E.C. observed	89.2	60.1	7.4	11.8

The B.E.C. of a clay might be estimated by adding up the exchangeable bases as determined by chemical analyses of the original clays, as in Table 23, which includes a good Wyoming, swelling bentonite, Volclay; a nonswelling, calcium bentonite from Panther Creek, Missouri; a kaolinitic plastic clay from Ohio; and a Kentucky ball clay. A ball clay acquires its plasticity from being carried in streams containing organic matter and salts and then being redeposited. Their chemical analyses are often intermediate between those of kaolin and montmorillonite. Many are kaolinitic but as so often happens, the different clay minerals are frequently mixed in the same deposit.

Base exchange is slow, requiring several days for substantial completion. On the other hand, one has to guard against altering the composition of the lattice itself by too prolonged dialysis or by treatment with too concentrated a solution for too long a time. Excessive grinding of clay renders most of its constituents soluble or replaceable.[16] In soil science it may be important to determine each of the ions replaced. In that case the clay is first extracted with water and then agitated for several days with a 1 N solution of ammonium chloride or ammonium acetate to replace the other ions with ammonium ion. After thorough and repeated treatment, the sample is analyzed for all possible cations. If only the total base exchange capacity is desired, a recommended method [17] is to leach the clay with 0.05 N hydrochloric acid for two days and let it settle and decant, the procedure being repeated at least four times. The product is finally washed free from excess hydrochloric acid and dried at 110° C. The B.E.C. is then determined by weighing out 0.2–0.3 gm of the residue, adding 100 ml of 0.01 N potassium hydroxide and shaking for three days. The sample is then back titrated with 0.01 N hydrochloric acid to the colorless range with phenolphthalein. Typical titration curves are shown in Figure 26·8.

Prolonged electrodialysis likewise leaves a hydrogen clay. Treatment with ion exchange resins is much quicker. The titration of a hydrogen clay may be carried out using a glass electrode. The B.E.C. can often be increased by preliminary treatment with twice normal hydrochloric acid. For example, that of the Panther Creek bentonite in the preceding table is raised from 60 to 105. The various ions differ in ease of replacement, lithium and sodium being most easily replaced, calcium, barium and hydrogen, least. It may well be an oversimplification of the subject to assume that calcium ions cannot exchange with elements in the layer structure itself and again that the magnesium, aluminum, and iron present are to be taken as being entirely in the layer structure and not, unless purposely added, among the external
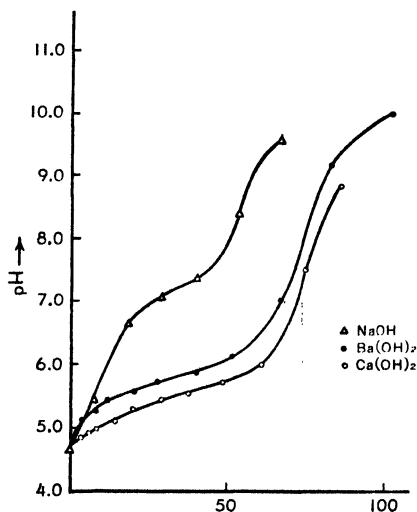
FIGURE 26·8 Typical titration curves of a bentonite. Abscissas show milliequivalents of base per 100 gm clay, ordinates show the pH. Taken from Mitra, R. P., *Bulletin* 4, Indian Soc. of Soil Science (1942), p. 145.

exchangeable ions. Ease of replacement has partly to do with the radius of the ion, which determines the ease with which the ion can fit into the hexagonal net of the external oxygen atoms. Even the difference in radius of anhydrous sodium and potassium atoms is significant in this connection. Ease of replacement is also dependent upon ionic hydration; compare Figure 26·2.

Puri and Uppal have emphasized that the B.E.C. of clay has a definite meaning only if the titration refers to a definite pH well on the alkaline side,[18] and they recommend use of the actual titration

curve. The titration curves with dilute sodium hydroxide resemble those of a weak monobasic acid, such as acetic acid, but the amount combined increases with the concentration of sodium hydroxide and also with the presence of salts, thus differing from the ordinary neutralization behavior of a soluble monobasic acid. Likewise, as was seen in Figure 26·8, the titration curves with different bases do not coincide. Each colloidal particle of bentonite can have thousands of charges upon it.[19]

Clays show very strong sorptive behavior, not merely for cations but for dyes, organic bases, and all sorts of other organic materials. The sorption is accompanied by increase in the long x-ray spacing with corresponding expansion of the spaces between the layers shown in Figure 26·2. Sorption of organic dyes takes place by two distinct complementary mechanisms, an irreversible ion exchange adsorption and a reversible physical adsorption. When basic dyes are sorbed on clays, a certain fraction of the sorbed amount is held irreversibly through base exchange, but as more dye is added, more is physically adsorbed onto the exchange-saturated clay.[20]

Marshall and Krinbill [21] refer to colloidal particles of clay as colloidal electrolytes on account of their charges and conductivity. A hydrogen clay paste has an appreciable conductivity and a pH between 2 and 3.5; whereas, if it is allowed to settle, the supernatant water is approximately neutral and exhibits very little conductivity. The conductivity curve of a hydrogen bentonite with change of concentration [22] has the typical form of a colloidal electrolyte such as soap, discussed in Chapter 17.

Base exchange has been employed in the construction of dams and artificial lakes to prevent leakage, as for example the lagoon on Treasure Island in San Francisco Bay. This is done by converting the clay into sodium clay, which has a very low permeability to water.

Base Exchange in Soils

Base exchange of soils was discovered 100 years ago by Thompson [23] and studied more particularly by Way.[24] The latter found that a solution of potassium chloride when passed through a bed of soil released calcium and sodium equivalent to the potassium sorbed.

The tenacious sorption of potassium ion is of the greatest importance in the fertility of soils since plants especially require potassium. The roots acquire it and other ions by direct base exchanges with the soil particles [25] and with soil humus, releasing unneeded hydrogen ions.

Thus the soil both stores and renders available the nutrient ions. The treatment of soils with calcium serves to make available the potassium in the soil. Tropical soils in rainy regions have much of the replaceable ions washed out by water, leaving acid or hydrogen clays. Soils of arid regions usually contain far too much exchangeable sodium, and normal agriculture depends upon restoring a suitable ratio of calcium to sodium. Calcium carbonate or sulfate is used; the gypsum, being more soluble, gives a quicker result. Sulfur alone may be used if the soil already contains calcium carbonate which will react to form gypsum.

Most anions are very weakly sorbed by clay. Hence nitrate fertilizer is readily carried away in drainage water and should be applied only when needed.

Normal field soils are aerogels with nearly half the pore space occupied by air and water between the secondary aggregates of the clay and soil particles. Soil and clay specialists in general assume that the clay particles are held in place by oriented molecules of water in the ionic layer between them; see Chapter 4. This view has been supported by the observations of Hauser and le Beau,[26] who found that monodisperse bentonitic suspensions of particles ranging from 200–960 Å apparent diameter exhibit a yield value when the particles are as much as 5.5 times their diameter from each other.

The Swelling of Clays

All clays swell in water to a certain degree. Hectorite is the most outstanding in this respect. When treated with water at room temperature it swells to a paste which looks exactly like boiled tapioca; further dilution gives a permanent suspension. There are different tests of swelling as well as different mechanisms of this phenomenon. One test is very gradually to add two grams of clay to 100 cc of distilled water and note the volume of the swollen mass settled at the bottom. If the true density of the clay is 2.5, its initial volume is 0.8 cc; but when powdered in air it occupies about 1.8 cc. A wide variety of clays so tested gave volumes ranging all the way from 2 cc to 45 cc. Kaolinitic clays range from 2–5 cc. The illites tested gave 5–11 cc, beidellite 6 cc, the montmorillonites from 4 cc upwards with no sharp demarkation between the "nonswelling" bentonites (5–15 cc) and the swelling bentonites that range all the way upwards to 45 cc. Only 1 clay out of 68 clays of all sorts tested gave instead only a suspension. There is no direct relation between swelling and the fineness of the particles.

For example, the dominance of surface reactions is illustrated by a bentonitic clay which showed 28% of particles one micron or less in water, but none in 1% sodium oxalate, yet swells to only 8 cc in water, whereas in sodium oxalate it swells to 24 cc. This type of swelling is related to base exchange but is by no means proportional to it. A sample of Wyoming "swelling" bentonite and another of Missouri "nonswelling" bentonite on being compared after base exchange each swelled in the following increasing order: [27]

$$Mg^{++}; \; Ca^{++}; \; H^+; \; K^+; \; NH_2^+; \; Na^+$$
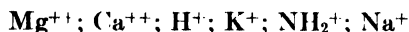
Even after a nonswelling bentonite had been converted by base exchange treatment to a higher B.E.C. value than a swelling sodium bentonite, it still did not swell quite as much as the original swelling bentonite.

It is clearly evident that this kind of swelling is closely related to sedimentation volume and to formation of a secondary structural aggregate of a brush heap nature. It is in sharp contrast to the primary swelling of the clay particles themselves as measured by x rays.[28] The maximum thickness of the water layer in a calcium or hydrogen montmorillonite is 10 Å and that in a sodium montmorillonite 21 Å, which is not even twice the volume of an air-dry clay.

Hence several mechanisms of swelling of clay must be simultaneously present:

First is the swelling of the planar water up to its limit of 10 or 20 Å.

Second is the necessarily different mechanism reported by Hauser and le Beau [29] and by Langmuir,[30] in which according to the ultramicroscopic evidence, clay particles are held far apart by the repulsive effect of the diffuse, ionic, double layer extending out from each of their surfaces. Thus, even a 0.5% suspension of bentonite particles of equivalent particle diameter below 140 Å exhibits a yield value, and a 1.4% system can set to a nearly transparent jelly. These particles are monoplates whose actual diameters are enormously greater than their thickness, which comprises only 10 Å of alumina and silica.

Obviously this is something different from the electric double layer in the sheets of planar water which remain only 10–20 Å thick within the crystal.

The third mechanism, which is of the greatest practical importance, though less amenable to theoretical treatment, is the structural viscosity caused by the ramifying aggregation which is common to most colloidal particles and which is exhibited in the swelling specimens first quoted. In more concentrated gels loose aggregates can be une-

quivocally demonstrated because protective colloids such as sodium
tannate lower the viscosity so greatly.

The effect of the ionic double layer at large distances, such as many
hundreds of Ångström units, has not been universally accepted, and
it is pointed out that it might be physically possible for the exceedingly
thin, disc-like particles of bentonite even in a very dilute dispersion to
be sufficiently in contact to build up a scaffolding structure.

After the discussion of base exchange it is not surprising to find
that the swelling of clay and its power of suspending other materials
require time to develop and may even pass through an optimum and
thereafter diminish in long time experiments.

Flocculation and Deflocculation of Clay

We have seen that clays will form suspensions in water, acquiring
a stabilizing charge, often by hydrolysis or by base exchange with any
salts present. Usually the charge on clay in water is negative. Kaolinitic
clay is positively charged only in very strongly acid solution. In general,
the addition of salts of strongly sorbed positive ions will serve to neu-
tralize the negative charge on the clay and bring about flocculation.
As would be expected, the trivalent cations are more effective than the
divalent, and these in turn than the monovalent cations, with the ex-
ception of hydrogen ion. Conversely, sorbed anions stabilize the sol
in the following decreasing order:

$$OH^-, SO_4^=, NO_3^-, Cl^-$$
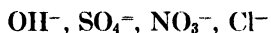
These rules are complicated by interactions with the materials already
in the clay particles.

In coagulation the particles clump together and settle out as flocs.
Frequently a smaller addition of electrolyte may result in formation
of a gel, that is, an aggregate that enmeshes the water.

As has been pointed out by Ford, Loomis, and Fidiam,[31] even if
the surface of a 3-layer clay such as hydrogen bentonite is negative,
broken edges exposing aluminum ions will be positive, whereas the
other edges will be negative. This increases the tendency of ramifying
aggregates to form through mere contact of edges and corner. Even at
the isoelectric point the observed moderate conductivity of clays is
significant, because the charges are still there but consist of an equal
number of positive and negative charges on the same particle, each
with its localized ionic atmosphere, as in the Debye-Hückel ion atmos-
phere around every ion in solutions of electrolytes.

The Rheology of Clays [32]

The addition of a small amount of water serves to turn a plastic thixotropic mass of completely swollen clay into a readily flowing slurry.

The viscous relations of clays are exceedingly complex and in summarizing them we will first follow the description of Lewis, Squires, and Thompson,[33] who describe the phenomena observed in a viscometer of a rotating concentric cylinder type such as the Stoermer. In it, the viscosity is proportional to the torque and inversely proportional to the rate of rotation. The first point to note is that a suspension of clay has a yield value. This is shown in Figure 26·9. Up to a certain torque there is almost no movement. Beyond that the rate of rotation is
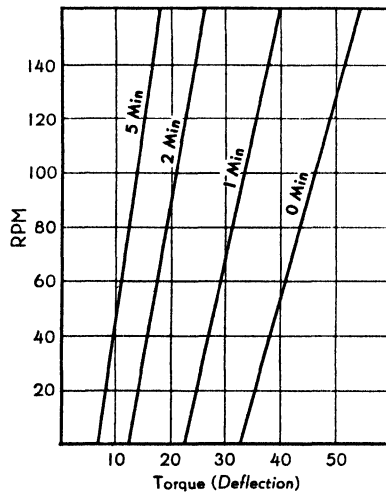
FIGURE 26·9 Effect of high-speed stirring on viscosity behavior of a drilling mud. After Lewis, Squires, and Thompson, Text Ref. No. 33.

practically proportional to the torque. Further, it is shown in Figure 26·9 that if the clay has not been stirred before testing, it has a high yield value, but that if it is stirred for 1, 2, and 5 minutes, respectively, the yield value is progressively lessened, and also the flow requires a smaller excess torque.

Further, the duration of stirring during the test affects both yield point and mobility. The latter is the slope of the straight line of flow against torque after the yield value has been overcome. In Figure 26·9 mobility is predominantly determined by the volume per cent of the

suspension and hence is relatively little affected by the large changes in yield value. This is shown in Figure 26·10. Thus thixotropy and yield values connote the presence of aggregates of clay particles with loose packing.

These phenomena are likewise dependent upon the type of suspending liquid, as is shown in Figure 26·11 for benzene, alcohol, water, and concentrated hydrochloric acid. Likewise these relations are greatly affected by the presence of peptizing agents, such as tannin, and by pH, as is shown in Figure 26·12. The mobility in alcohol is practically the same as that in water although the yield point is 50% greater. Hence there is no direct connection between yield value and mobility.
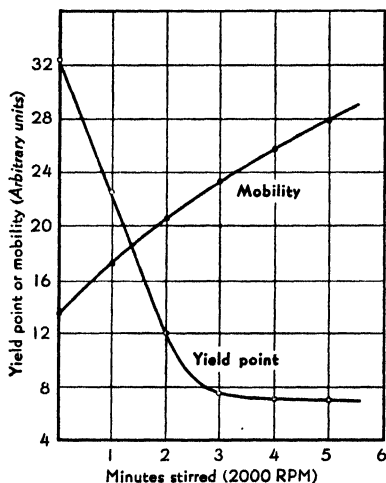
FIGURE 26·10 Effect of high-speed stirring on yield point and mobility of a drilling mud. Taken from Lewis, Squires, and Thompson, Text Ref. No. 33.
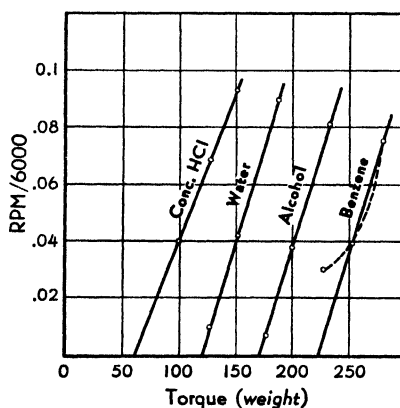
FIGURE 26·11 Effect of type of suspending liquid on clay suspensions. Concentration of clay equals 60 gm per 100 cc of liquid. Taken from Lewis, Squires, and Thompson, Text Ref. No. 33.

The hypothesis of aggregates has been supported by the following observations: (1) The amount of additive required for a given reduction in viscosity is greater for fine particles than for coarse particles as if the surface were coated to prevent aggregation. (2) After standing, the viscosity often increases again, ions diffusing into the interior of the particles and no longer affecting the aggregation and sometimes even splitting the clay into more platelets with increased charge and possibility of contact. (3) The relative amounts of sodium tannate and sodium pyrophosphate required for a given reduction of viscosity

are as 1 : 5, but the area which could be covered by sorbed tannate ion is about 5 times larger than that for a pyrophosphate ion.

A yield value becomes noticeable in bentonite at concentrations lower than that at which thixotropy first appears. It is ascribed [34] to structure in the clay suspension. Coacervation may occur either before or after thixotropy becomes noticeable.

Broughton and Hand [35] point out that with a kaolinitic ball clay the yield value increased with temperature, being 22.6 dynes per square centimeter at 25° C and 32.0 at 50° C. This they point out contradicts the attraction-repulsion force theory of Hamaker based on a long range attraction across the diffuse double layers. Likewise they found
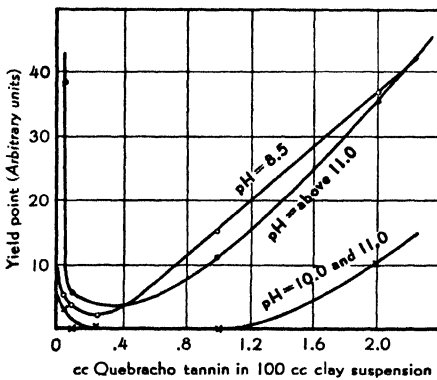
FIGURE 26·12 Effect of peptizing agents (Quebracho tannin) on yield point of a drilling mud at various pH values. Taken from Lewis, Squires, and Thompson, Text Ref. No. 33.
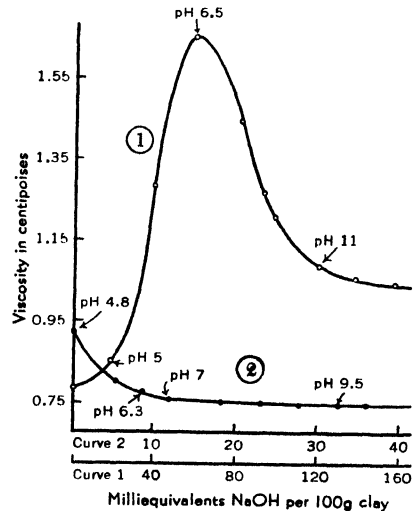
FIGURE 26·13 Comparison of the viscosity of montmorillonitic hydro clays and that of kaolinitic hydrogen clays at various pH values. Concentration of clay not given. After Mitra and Mukherjee, Text Ref. No. 36.

no thixotropy in clay suspensions containing sodium hydroxide, but very pronounced thixotropy with barium hydroxide added to bentonite from pH 4 upwards.

Mukherjee and Mitra [36] have pointed out a striking difference between kaolinitic and montmorillonitic clays. The viscosity of a kaolinitic clay is almost unaffected by further additions of sodium hydroxide after the initial decrease. In contrast, that of a bentonitic clay passes through a high narrow maximum at about pH 6.5, as is shown in Figure 26·13.

The Practical Testing of Clay

The preceding discussion has incidentally mentioned some of the tests such as the flow characteristics, swelling, and gelation that contribute to the evaluation of the useful properties of a clay. A complete syllabus of clay testing has been issued by the Bureau of Mines.[37] The tests range from the simple but informative Atterberg test for plastic properties, through chemical analyses and drying and firing tests for ceramics, to slaking, bleaching, and activation for other uses.

Uses of Clay

The primary use of clay is in agriculture for it is principally the clay in soils that retains water and minerals necessary to plants. The planar water does not freeze; even that held in capillary spaces has a lowered freezing point of $-5°$ for a pore radius of 250 Å [38] and is held under a tension of about 57 atmospheres. The tension is inversely proportional to the pore diameter and would therefore be 1000 atmospheres for pores of 14.2 Å.

The oldest use of clay is for pottery, and the finest china is made from kaolinite. The clay has to be plastic and may require weathering of organic and other constituents for prolonged periods to produce sufficient plasticity. The moist clay is plastic, retaining its shape when deformed, the platelike particles adhering as the water is pressed from between them. The shape is then made permanent by heating to drive off water and convert the clay into other chemical species. At one extreme the product is brick or tile and at another refractory or fire clays.

The clays for ceramic wares are frequently deflocculated by alkaline solutions such as sodium silicate and sodium carbonate, which permit the casting of slips in molds. Bentonitic clays are used as suspending agents in enamels.

An intriguing use is that of the alsifilms developed by Hauser and le Beau.[39] On drying, gels of bentonite form coherent, translucent films which can be made nonswellable by ion exchange. These are excellent insulators, resembling mica, and are oil and heat resistant and capable of being plasticized and made flexible.

Clay is important for its sorptive properties in the bleaching and refining of many materials, such as vegetable oils, as is evidenced by the use of the term "Fuller's Earth." Many such clays are bentonitic although some are composed of attapulgite. For bleaching, a bentonite

clay is activated with acid which extracts alumina and ferric oxide, after which the clays are dried at an elevated temperature. For catalytic application they are heated between 400 and 870° C. It is therefore evident that the catalytic activity is no longer ascribable to montmorillonite as such. Such activated clays have found a very important application in the petroleum industry in the production of lubricating oils by polymerization.

The original alumina silicates often give deep colors with aromatic amines and Vitamin A.[40]

Clays are used for removing organic matter in chromatography and in cosmetics and soaps. Clay is used as a coating on paper and also in treating effluent from paper manufacture. A very large number of uses have been listed in the reports by Spence and by Davis and Vacher already cited.

Marshall [41] has developed clay membranes for determining the concentration of sodium and other ions, analogous to the use of the glass electrode in measuring pH of solutions.

The use of clay in drilling muds should be briefly discussed.[42]

During the last few decades bore holes for oil and gas wells have been drilled with a rotary drill attached to the bottom of a drill pipe. Down the pipe is passed a drilling fluid or mud to the bottom of the hole, whence it passes up again to the surface through the annular space between the drill pipe and the wall of the hole and is then recirculated by means of pumps. The volume of mud pumped commonly amounts to several hundred gallons per minute under pressures ranging from a few hundred to a few thousand pounds per square inch.

The primary function of a drilling mud is to transport cuttings to the surface, but if the circulation is temporarily stopped, the mud must be thixotropic and capable of holding the cuttings in suspension. The high hydrostatic pressure of the drilling mud prevents other fluids, such as oil, gas, water, and brine, from coming into the hole, but the mud itself would be lost by going into the formation if it were not for its wall-building property, that is, the building up of a more or less impervious filter cake on the wall of the bore hole. This also strengthens the wall and minimizes caving, heaving, and sloughing. Likewise the drilling mud cools and lubricates the drill stem and bit, carries part of its weight, and brings samples of oil and other materials encountered to the surface. This is a remarkable combination of properties which has enabled holes to be drilled more than three miles deep even where porous formations, salt and gypsum beds, brine strata, and other difficult formations have been met.

Although drilling mud must be sufficiently thixotropic to suspend the cuttings of the bore hole, its yield value under agitation must be just sufficiently low that when the mud flows horizontally in ditches at the surface, gas may be released and the cuttings drop out. Foam breakers are often added to help release the gas. Also, the filter cake on the wall of the bore hole must remain a very thin layer so as not to obstruct the hole. The best clays for this purpose are the swelling, Wyoming bentonites which combine all the properties enumerated, although other clays may possess them in varying degree. However, even better wall building, without the other properties, is obtained by boiled starch, which has the advantage of not being coagulated by concentrated salt solutions sometimes met in drilling operations.

The properties of a drilling mud are usually controlled by the addition of such materials as sodium tannate or various phosphates. Sodium silicate may be added to counteract coagulation by salts, or chemicals may be employed to remove or precipitate coagulating ions. Attapulgite is much less affected than bentonite by salts. To increase the hydrostatic pressure, barytes is added to increase the density of the mud and it is kept in suspension by the mud. A typical drilling mud would consist of 65–98% by volume of water, 2–30% by volume of clay (the lower figure being for the highest swelling clay), 0–35% of finely powdered barytes, and 0–10% by volume of materials coming in from the drilling operation. Occasionally an oil-base drilling mud is used. Typical oil soluble deflocculants are oleic acid; nitrogen bases, such as hexyldecylamine, xylidine, piccoline; and alkyl sulfonates and sulfates. Emulsions containing soaps have recently met with success.

It is evident that a careful study of each of the factors in the rheology of clays is of importance in the design and control of drilling muds.

Similarities and Contrasts between Different Types of Colloids

Since we are now bringing to an end our review of some of the more prominent phenomena met with in the commoner colloidal materials, we may pause to recall some of the typical and some of the contrasting behavior of important colloids. First we may quote from Lewis, Squires, and Thompson [43]

> . . . there are striking parallelisms between clay suspensions and gelatin solutions. Both are colloidal. Both exhibit Brownian movement. Both are highly viscous. Both set under proper conditions to strong and stable gels. The colloidal properties of both are profoundly influenced by the hydrogen-ion concentration (pH),

showing in each case marked maxima and minima at definite pH values. The two are similarly affected by electrolytes, though in different degrees. Despite these parallelisms, the differences are fundamental. The outstanding distinction between the two types is found in the fact that in low concentrations, materials of the clay type have high and those of the gelatin type may have low mobility, relative to the fluidity of the solvent. Again, while gelatin is water soluble, it is highly resistant to most organic solvents. It is completely unaffected by hydrocarbon liquids, and is precipitated from aqueous solutions by alcohol and similar organic solvents miscible in water. The addition of small amounts of these precipitants results in a marked decrease in viscosity of the gelatin solution. Clay, on the other hand, is plasticized not only by the water but by these very organic solvents that precipitate gelatin. Indeed, using yield point as a measure of plasticization, the organic liquids are more effective than water itself. Furthermore, it is possible to reduce the yield point of a clay suspension to a negligible value (by treatment with tannic acid at suitable pH), but such treatment changes the mobility little, and the viscosity of the suspension is influenced but slightly by the method of breaking the yield point by other conditions, so long as the yield point is low. Gelatin solutions also show a yield point under proper conditions and this yield point can be reduced to a negligible value; e.g., either by agitation or rise in temperature. However, here the analogy ends, because there is a large change in mobility with drop in yield point and the viscosities of solutions of zero yield point are profoundly influenced by variation in conditions; e.g., pH. Solutions of the type of gelatin in water, and rubber in suitable organic solvents, when not too concentrated, show a sharp gelation temperature, below which they will set to a gel but above which they do not. Such solutions in general show no yield point at temperatures above their gel point, but at these temperatures their viscosities can be profoundly influenced by addition agents and at these temperatures if their viscosity-temperature curve is extrapolated downward it approaches an infinite value of the viscosity asymptotically at the gel point; or, what is equivalent to the same thing, the fluidity becomes zero at that temperature.

A gelatin has analogs in such substances as agar. In contrast with gelatin, which becomes less viscous with time, are the types represented by *silicic acid*. These sols steadily increase in viscosity until they become a glass. This occurs even faster at higher temperatures and occurs even with solutions of sodium silicates. Freshly formed silicic acid is in true solution. Its sols appear most stable in very dilute acid at the isoelectric point, where most colloids are least stable. A silicic acid sol is most unstable in dilute alkali, the opposite to what one would expect from the chemical composition. As in so many cases, salts follow the lyotropic series and again, as so often happens, the effects depend greatly, as regards rate and direction, upon the presence of hydrogen and hydroxyl ion. For example, addition of 1 N potassium chloride gelatinizes silicic acid in one day, but in the presence of a few drops of alkali the gelatinization is instantaneous. If any strong acid is present, the salt has no effect. There is evidence [44] that in silicic acid gels there are both fibrous aggregates and particles which are held apart by secondary forces. This contrasts with ferric hydroxide, which

is not fibrous but platelike, and can form Schiller layers held apart by long range effects.

Analogs of silicic acid are found in stannic, tannic, molybdic, and tungstic acids and in aluminum hydroxide. In many of its properties starch is similar, but here hydrogen ion coagulates the starch and hydroxyl ion does not.

These behaviors can again be contrasted with that of albumin, which does not gelatinize but which can be salted out reversibly by rather concentrated solutions of the salts of alkali metals. The salting out turns to irreversible coagulation in sufficiently acid solution. Irreversible coagulation is also brought about by heat, which liquefies a gelatin jelly and scarcely affects a clay. However, like clay, proteins are readily coagulated by salts of the heavy metals.

It is clearly seen that the colloidal behavior of these materials is almost indefinitely complex but is subject to definite factors and may be controlled in any desired direction for practical purposes.

REFERENCES

Text References

1. RIES, H., *Clays, Their Occurrence, Properties and Uses*, 3rd ed. John Wiley & Sons, Inc., New York (1927)
2. KEEN, B. A., *The Physical Properties of the Soil*, Longmans, Green and Co., London and New York (1931)
   BAVER, L. D., *Soil Physics*, John Wiley & Sons, New York (1940)
   DALLA VALLE, J. M., *Micromeritics, the Technology of Fine Particles*, Pitman Publishing Corp., New York and Chicago (1943)
   PIPER, C. S., *Soil and Plant Analysis*, Interscience Publishers, Inc., New York (1944)
   A whole series of voluminous but uncritical books on clay, from the standpoint of their practical properties and their uses in ceramics, refractory materials, and brick-making have been written by A. B. Searle
   MARSHALL, C. E., *Colloid Chemistry of the Silicate Minerals*, Academic Press, New York (1949)
3. KRUMBEIN, W. C., *J. Sediment. Petrol.*, 2, 89–124 (1932)
4. HAUSER, E. A. and REED, C. E., *J. Phys. Chem.*, 40, 1169 (1936)
5. GRIM, R. E. and BRAY, R. H., *J. Am. Ceramic Soc.*, 19, 307 (1936)
6. BANCROFT, W. D., *Applied Colloid Chemistry*, McGraw-Hill Book Co., New York (1932)
7. SPENCE, H. S., *Bentonite*, Department of Mines, Ottawa, Canada (1924)
   ROSS, C. S. and SHANNON, E. V., "The Minerals of Bentonite and Related Clays and Their Physical Properties," *J. Am. Ceramic Soc.*, 9, 77–96 (1926)
   ROSS, C. S. and KERR, P. F., *The Kaolin Minerals*, U.S. Geological Survey, Professional Paper 165 (1930), p. 151
   GRIM, R. E. and BRAY, R. H., *The Mineral Constitution of Various Ceramic Clays*, U.S. Geological Survey, Professional Paper 165 (1930)
   BRAGG, W. L., *Atomic Structure of Minerals*, Cornell Univ. Press, Ithaca (1937)

GRIM, R. E., *Properties of Clay*, Symposium on Recent Marine Sediments, American Association of Petroleum Geologists (1939)

HENDRICKS, S. B. and ALEXANDER, L. T., "Minerals Present in Soil Colloids," *Soil Sci.*, 48, 257 (1939)

DAVIS, C. W., VACHER, H. C. and CONLEY, J. E., *Bentonite, Its Properties, Mining, Preparation and Utilization*, Bureau of Mines Tech. Paper 609 (1940)

GRIM, R. E., "Modern Concepts of Clay Minerals," *J. Geol.*, 50, 225–273 (1942)

ROSS, C. S. and HENDRICKS, S. B., *Minerals of the Montmorillonite Groups*, U.S. Geological Survey, Professional Paper 205B (1943-4), pp. 23–70

BAGCHI, S. N., *Indian J. Phys.*, 29, 53–80 (1946)

8. HAUSER, E. A., *J. Am. Ceramic Soc.*, 24, 179 (1941)

9. JACKSON, M. L., Ph.D. Thesis, Univ. of Wisconsin, 1939

10. MACEWAN, D., *Nature*, 154, 577 (1944); 157, 159 (1946); *J. Soc. Chem. Ind.*, 65, 298 (1946)

11. JACKSON, M. L. and HELLMAN, N. N., *Soil Sci. Soc. Proc.*, 6, 133 (1941); ALDRICH, D. G., HELLMAN, N. N. and JACKSON, M. L., *Soil Sci.*, 57, 215 (1944); BRADLEY, W. F., *J. Am. Chem. Soc.*, 67, 975 (1945); WHITE, J. L. and JACKSON, M. L., *Soil Sci. Soc. Proc.*, 11, 150 (1946); MACEWAN, D., *Trans. Farad. Soc.*, 44, 349 (1948)

12. SPEIL, S., BERKELHAMER, L. H., PASK, J. A. and DAVIES, B., *Tech. Paper # 664*, Bureau of Mines, U.S. Dept. of Interior, 1945; GRIM, R. E. and ROWLAND, R. A., *Amer. Mineral.*, 27, 746–761 (1942); 801–818; *J. Am. Ceramic Soc.*, 27, 5 (1944)

13. PREBUS, A. F., "The Electron Microscope," in J. ALEXANDER, *Colloid Chemistry*, Vol. V, Reinhold Publishing Corp. (1944), p. 152

14. HENDRICKS, S. B., *Ind. Eng. Chem.*, 37, 625 (1945)

15. THIESSEN, P. A., *Z. Electrochem.*, 48, 675 (1942)

16. JACKSON, M. L., *loc. cit.*

17. COFFMAN, P. M., M.A. Thesis, Stanford University, 1946

18. PURI, A. N. and UPPAL, H. L., *Soil Sci.*, 47, 245 (1939)

19. BRADFIELD, R., in J. ALEXANDER, *Colloid Chemistry*, Vol. III, Chemical Catalog Co., Inc., New York (1931), p. 578

20. PLESCH, P. H. and ROBERTSON, R. H. S., *Nature*, 161, 1020 (1948)

21. MARSHALL, C. E. and KRINBILL, C. A., *J. Phys. Chem.*, 46, 1077 (1942)

22. MUKHERJEE, J. N. and SEN GUPTA, N. C., *Nature*, 145, 971 (1940)

23. THOMPSON, H. S., *J. Roy. Agr. Soc.* (England), 11, 68 (1850)

24. WAY, J. T., *J. Roy. Agr. Soc.* (England), 11, 313 (1850)

25. JENNY, H. and OVERSTREET, R., *Soil Sci.*, 47, 257 (1939); 55, 257 (1943)

26. HAUSER, E. A. and LE BEAU, D. S., *J. Phys. Chem.*, 45, 54 (1941)

27. COFFMAN, P. M., *loc. cit.*

28. HOFMANN, U. and BILKE, W., *Kolloid-Z.*, 77, 238 (1936)

29. HAUSER, E. A. and LE BEAU, D. S., *J. Phys. Chem.*, 42, 961, 1031 (1938)

30. LANGMUIR, I., *J. Chem. Phys.*, 6, 873 (1938)

31. FORD, T. F., LOOMIS, A. G. and FIDIAM, J. F., *J. Phys. Chem.*, 44, 1 (1940)

32. HAUSER, E. A., *Chem. Rev.*, 37, 287–321 (1945)
MUKHERJEE, J. N. and SEN GUPTA, N. C., *loc. cit.*
GARRISON, A. D., *Tech. Publ. # 1027* of Am. Inst. Min. & Metall. Eng. (1939)
REED, C. E., *Tech. Publ. # 871* of Am. Inst. Min. & Metall. Eng. (1937)

33. LEWIS, W. K., SQUIRES, L. and THOMPSON, W. I., *Trans. Am. Inst. Min. & Metall. Eng.*, 114, 38 (1934)

34. MUKHERJEE and SEN GUPTA, *loc. cit.*

35. BROUGHTON, G. and HAND, R. S., *Nature*, 142, 255 (1938)

36. MUKHERJEE, J. N. and MITRA, R. P., *J. Coll. Sci.*, 1, 141, 157 (1946)

37. KLINEFELTER, T. A., O'MEARA, R. G., GOTTLIEB, S. and TRUESDALE, G. C., *Bull. # 451*, Bureau of Mines, Washington (1943)

38. MCBAIN, J. W., *Sorption of Gases and Vapours by Solids*, Geo. Routledge and Sons, Ltd., London (1932)

39. HAUSER, E. A. and LE BEAU, D. S., *J. Phys. Chem.*, **43**, 1037 (1939)

40. WEIL-MALHERBE, H. and WEISS, J., *J. Chem. Soc.*, 2164 (1948)

41. MARSHALL, C. E., *J. Phys. Chem.*, **48**, 67 (1944)

42. STERN, A. G., *Role of Clay in Oil Well Drilling Fluids*, U.S. Bureau of Mines, R. I. 3556, February, 1941

    LOOMIS, A. G., *Tech. Publ.* # 1201, Am. Inst. of Min. and Metall. Eng. (1939)

    EVANS, P. and REID, A., *Trans. Min. & Geol. Inst.* (India), **32**, 1–261 (1936)

    LARSEN, D. C. H. in J. ALEXANDER, *Colloid Chemistry*, Vol. VI, Reinhold Publishing Corp. (1946), pp. 509–530

43. LEWIS, SQUIRES, and THOMPSON, *loc. cit.*, pp. 50–51

44. TOURKEY, A. R., *J. Soc. Chem. Ind.*, **61**, 254 (1942)

The behavior of colloidal dispersions in gases such as air is quite different from that which we have been reviewing in the presence of a solvent. Aerosols are always unstable and tend to aggregate and then precipitate. Although they are of rather general importance, most of our knowledge about them arises from what has been revealed by results obtained in the course of the studies of the civilized nations in preparation for war.

The two chief classes of aerosols are (a) those whose particles are nonvolatile, whether solids or liquids, such as dusts, fumes, smokes, smoke screens and fogs, and bactericidal mists and (b) those whose particles are volatile and usually aqueous, such as mists, clouds, and fogs.

The general behavior of these systems can be indicated by describing a vivid demonstration given to the author by Dr. Whytlaw-Gray. An electrical spark was passed between two silver wires for a few seconds in a closed chamber, one meter cube. Samples of air were then withdrawn from the chamber and passed through an ultramicroscope with arcs on two sides to give greater intensity of illumination. At first, however, nothing was visible. Soon particles began to appear, rapidly increasing in number and showing violent Brownian movement. The particles became brighter and moved more slowly until some were brilliantly visible. Evidently at first the vapor of silver had condensed into particles far too small to be revealed in the ultramicroscope; and since silver particles are not volatile, they must have been aggregating until the secondary particles were large enough to be seen. Application of an electric field showed particles which moved toward one electrode, others moving toward the other electrode, and therefore

of opposite charge, and some that appeared unaffected. After an interval some glass microscope slides that had been lying in the bottom of the original chamber were taken out and examined under the microscope with intense sideways illumination. It was seen that the silver dust had settled upon the glass in a remarkable way. Frequently this took the form of long flexible chains one end of which was firmly attached to the glass while the rest waved about in the air. The chains were not visibly continuous but more resembled necklaces, like beads upon an invisible thread. Some of the chains were branched, and Brownian movement could be observed within and along them.

A description of these "precipitates" may be quoted from Whytlaw-Gray and Patterson, *Smoke; A Study of Aerial Disperse Systems,* p. 79:

> Evidence which shows that the units in these complexes are not actually in contact is afforded by some observations we have made ourselves on the branching tree-like chains of zinc oxide and cadmium oxide smoke particles which collect on the electrodes in an electric field. These can be made to rotate continuously in a gentle stream of air.
>
> In some the whole tree-like structure will spin round whilst the trunk remains stationary and in others the separate branches only will exhibit this remarkable movement.

The remainder of the phenomena involved in aerial dispersions are better described in the more systematic treatment below, but the above serves to emphasize the instability of an aerosol and its tendency toward aggregation, often in a very loose or directed pattern within which Brownian movement may occur. The aggregation of zinc oxide smoke, as seen in the ordinary ultramicroscope, is illustrated in Figure 27·1. The form of the actual particles is shown by the electron microscope to be acicular, as is seen in Figure 27·2. In Figure 27·2 the particles are about 450 Å thick and 4500 Å long; they were formed by burning zinc in air.

Preparation of Aerosols

The chief method of forming aerosols is by the sudden condensation of highly supersaturated vapor. The following procedures have been developed for this purpose. (a) The material may be heated in a stream of air or other gas. (b) It may be dissolved in a very volatile propellent solvent such as the nontoxic substance Freon, CCl_2F_2, and suddenly released. This is the favorite method of producing an aerosol of the insecticide DDT; usually some oil is also dissolved in the Freon. (c) Materials may be heated in an electric arc. Platinum, silver, and
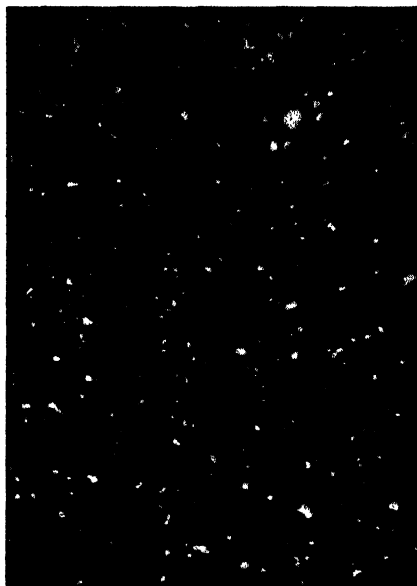
9 Minutes Exposure.

29 Minutes Exposure.

49 Minutes Exposure.
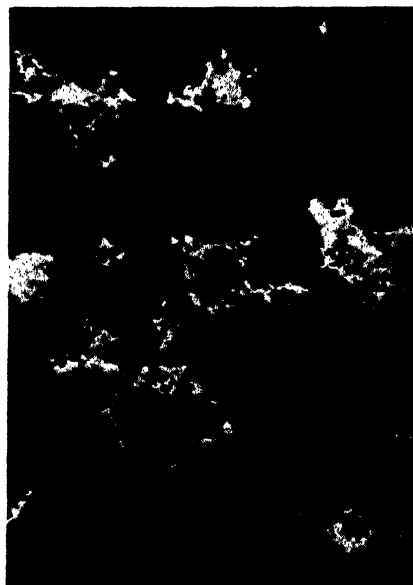
244 Minutes Exposure.

FIGURE 27·1 Aggregation of zinc oxide smoke as seen in the ordinary ultramicroscope. Taken from Whytlaw-Gray, Speakman, and Campbell, *Proc. Roy. Soc.*, 102A, 600 (1923).
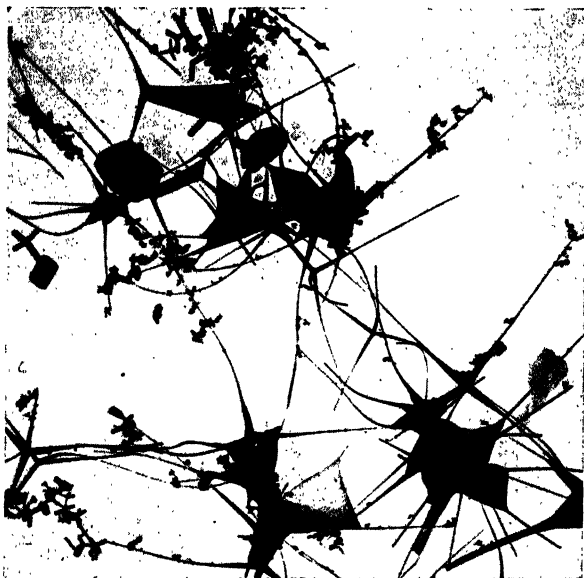
FIGURE 27·2 Particles of zinc oxide smoke as seen in the electron microscope. Magnification 15,000 ×. Courtesy Farrand Optical Company.

gold volatilize as such. The other metals form oxides. (d) Aerosols may be detonated with fulminate. (e) They may be produced by controlled combustion of, for example, a mixture of potassium chlorate and sugar. (f) They may be produced by chemical reaction of such substances as zinc ethyl, phosphorus, or magnesium or by the reaction between ammonia and hydrochloric acid. (g) Smoke screens or fogs of exceptional efficiency may be produced by releasing the heated material through a cone-shaped nozzle.[1] (h) Dusts and smokes and "smog" are produced as the result of many industrial operations and combustions. (i) In nature, cloud particles are produced by condensation of water vapor. (j) Ultrasonic dispersion of nonmetallic liquids such as benzene, toluene, water or oils is sometimes used. (k) A smoke screen may be produced by reaction of a stock mixture of grained aluminum (10.1%), zinc oxide, and hexachloroethane, producing a zinc chloride with all the carbon appearing as smoke. Less aluminum results in formation of carbon monoxide or even carbon dioxide instead of carbon.[2] (l) Infective aerosols arise from living plants (spores) and animals, dust, and scaling skin, etc.; see Figure 27·3. (m) Natural dust clouds arise from volcanic eruptions, such as that of Krakatoa, and from calamitous erosion of agricultural soil. (n) Sudden volatilization may produce an aerosol as in the case of anthracene, o-phthalic acid,

and paraffin wax but not in others, such as clean mercury or benzyl-benzoate. Dyes are volatilized to give colored smokes. (o) Fairly stable fogs are produced by atomizing hygroscopic materials, such as concentrated sulfuric acid or silicon chloride, which then pick up water vapor from the moist air. (p) Another exceedingly fine aerosol with highly electrified particles is obtained when a gas is passed over a heated metal.[3] These give rise to visible luminescence at a distance of one or two feet away from the source when they touch the walls of a silica tube. Their presence has been verified and their numbers have been measured although they are too fine to show in the ultramicro-scope. They have frequently been stated to be capable of promoting chemical reactions.

The formation of particles and, in particular, of liquid droplets is controlled by the formation and presence of nuclei. By providing condensation nuclei monodisperse clouds are obtainable.[4] The formation of colloidal particles on a nucleus is extensively discussed by Sved-berg.[5] Nuclei may be infinitesimal salt particles, as within a range of several hundred miles of an ocean, or often inland. Frequently they are electric charges or gas ions, as is familiar in the cloud chambers originated by C. T. R. Wilson and used in the study of charged particles, cosmic rays, and nuclear fission. According to Rodebush [6] ordinary water contains more or less permanent nuclei. Often they are the Langevin atmospheric ions of average radius 200–300 Å.[7] Sometimes they are hygroscopic particles such as ammonium chloride, sulfuric

FIGURE 27·3 Introduction of microorganisms into the air — a violent sneeze. 1/30,000th second exposure. Over 40,000 particles are shown. Courtesy Marshall W. Jennison.

acid, or nitrous acid. Smokes consisting of silver iodide produced by vaporization of silver iodide serve as nuclei for the condensation of water as ice crystals at temperatures below $-4°$ C.[8]

The smokes produced by an arc are highly electrified and have an unusual tendency to form chains on coagulation. Many smokes are not so electrified. Crystals that are pyroelectric such as aminoazo benzene also aggregate in long strings that appear somewhat discontinuous.[9]

Concentration and Number of Particles

An aerosol is usually very dilute and its concentration may be measured in milligrams per cubic meter. The particles are generally not particularly minute. They usually exceed 1000 Å in diameter. Fogs may reach a size of as much as a few tenths of a millimeter. The particles of a smog show most frequent values around 7 microns.[10] Langmuir has shown that the optimum size of particles in a screening fog to give maximum light scattering and produce an opaque effect is 0.31 micron.[11] The number of particles per cc of typical aerosols is very low, being of the order of 10^4 per cc as compared with a concentration of ions in pure water of 10^{11} per cc. Commercial dusts are usually between 0.2 and 2 microns.

The concentration of aerosols has been determined by a number of methods, such as the use of a very fine filter which is afterwards weighed on a microbalance. This is indeed the principle employed by gas masks to remove poisonous "particulate gases." The particles are not screened out by the use of interstices of smaller diameter than the particles. This is quite unnecessary and would make breathing through the gas mask difficult or impossible. Instead, a mat of loose fibers such as cellulose is employed, and the gas stream has to follow the relatively wide but sharp convolutions and turns in the interstices. This centrifuges out the heavier particles from the gas, and they stick to the surfaces of the materials in the filter pad.

In artificially produced aerosols the content is often known from the mass of the original materials. Many other special methods have been used, such as measurements of the absorption of light or the intensity of light scattering (R. C. Tolman, 1919), the depth of stain on a filter, burning an asbestos filter and determining carbon dioxide produced, determining ammonium chloride by Nessler reagent, the use of dust counters such as those devised by Owens or Aitken (both of which have been criticized), use of a microprojector for dust,[12] and

culturing of bactericidal aerosols after filtration. For determining visibility over an airport a device is used which responds solely to a tuned, intermittent light source, independent of all other light.

Green [13] has developed the cloud chamber of C. T. R. Wilson to produce expansions in the cell of an ultramicroscope. Under these conditions the particles act as nuclei, and the water droplets formed are large enough to be photographed and counted. This method is applicable to the very finest aerosols, even those quite invisible in an ultramicroscope. For coarse particles of the order of 1 micron Green has used a thermal deposition of particles on a pair of microscope cover glasses between which is placed an electrically heated wire. [14]

Some of the above methods also give the number of particles. Counts of these have been obtained with the ultramicroscope by counting particles collected on slide glasses or by the use of photoelectronic counting devices. [15] An optical mass concentration meter has been employed by La Mer and associates.

The Spontaneous Coagulation of Aerosols

Almost every encounter between a particle of an aerosol and a surface, or another particle, is inelastic and results in permanent adhesion (although a strong current of air can release the deposited particles). Von Smoluchowski has formulated the quantitative behavior of such coagulation. First, however, we may describe the observed experimental behavior. Figure 27·4 illustrates the way in which the number of particles per cc, N, varies for aerosols of stearic acid, resin, and magnesium oxide, respectively. The same data are plotted differently in Figure 27·5 using instead the reciprocal of N, which now represents the average space inhabited by a particle. The reciprocal of the number of particles per cc is in all cases nearly or quite linear when plotted against time according to the equation:

$$\frac{1}{N} = \frac{1}{N_0} + Kt$$

where N_0 is the count at zero time, and K is the constant for the given smoke. This equation is identical in form with that established by Rutherford in 1897 for the rate of combination of small ions. It holds also for the combination of small ions and nuclei, for large ions with each other, and also for the rate of disappearance of nuclei. The slope K depends upon the concentration or total weight of material in the aerosols of originally the same number concentration. The higher

concentrations coagulate somewhat more slowly. It follows that the larger the particles the more slowly they coagulate. This is completely at variance with the behavior of sols in liquids for which the coagulation constant does not depend upon the size of particle.
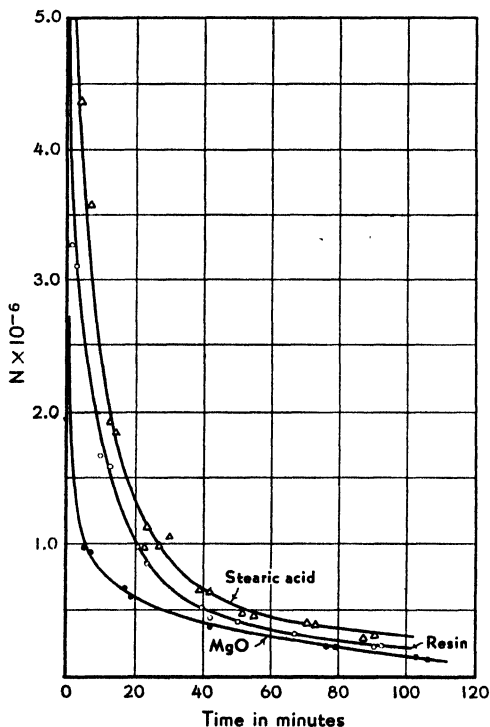
FIGURE 27·4 Number (N) of particles of three aerosols as designated, plotted against time. Taken from Whytlaw-Gray and Patterson, *Smoke* (see References).

Another significant observation is that the more homogeneous a given aerosol is, the more slowly it coagulates. There are two sources of disturbance in studying the coagulation of aerosols. The first is the loss of particles which strike any exposed surface, and the second is evaporation or volatilization or distillation from one particle to another. The theory of von Smoluchowski, which does not take into account these factors or electrical charges, has been adapted by Whytlaw-Gray [16] for application to aerosols. It is simplest for monodisperse aerosols. It has been shown to take a form equivalent to the empirical equation already mentioned in which the constant

$$K = \frac{2\,RT\,s}{3\,\eta\,N}\left(1 + A\,\frac{l}{r}\right)$$

where R is the gas constant, T the absolute temperature, η the viscosity of the medium, and s is the ratio of the radius of the sphere of influence around a spherical particle to the actual radius of the particle. In general, s is found to be equal to 2 within the limits of error, which means that 2 particles stick together whenever they are brought into contact by Brownian movement. The agreement with experiment is usually excellent. In the usually small correcting term, A is a constant, l the mean free path of the molecules of the gaseous medium, and r the radius of a particle.
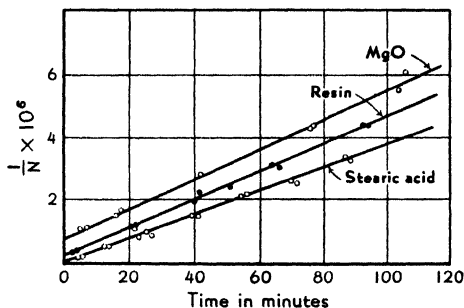
FIGURE 27·5 Reciprocal of N, plotted against time, for the same aerosols as are shown in Figure 27·4. Taken from Whytlaw-Gray and Patterson, *Smoke* (see References).

Polydisperse smokes coagulate differently. Small particles are very rapidly eliminated by collision with the larger ones. If then this smoke is rapidly diluted by the larger particles beginning to cluster, the aerosol remains fairly homogeneous. If the particles are not spheres but are irregular, the rate of coagulation is increased. Aerosols of stearic acid appear to be approximately spherical. Those of resin and of hygroscopic particles are liquid droplets, but chainlike structures are produced from the highly electrified smokes from an arc. However, chainlike aggregates are obtained with the uncharged particles from zinc ethyl. Hence they are due in such cases to the polarity of the oxide particles. The rate of coagulation is not much influenced by electrical charges. The treatment given has assumed a quiet medium; coagulation is greatly hastened by turbulent agitation.

The aggregates produced by coagulation are very porous. Their size may be determined by applying the Stokes-Cunningham law for rate of fall in air. In general, it is found that their density is only 1/11 to 1/6 of the density of the material of which the actual particles consist. Hence the aggregates are many times larger than they would appear to be by light scattering or by examination in the ultramicro-

scope. Frequently Brownian movement can be observed within the aggregates. Chainlike aggregates may be as much as 30 microns long, although only 500 Å wide.

The fact of aggregation has been confirmed by treatment with various liquids, such as water or oil, which produce very fine colloidal solutions, showing that the aggregates are readily separated into their finer particles. Whytlaw-Gray states that "they are not actually in contact but are kept apart by films of air thin enough for cohesional or molecular forces to be effective."

It is still a moot point in many cases as to whether or not there is such a thing as a definite primary particle in the aerosols. Presumably there is in the zinc oxide crystals that were revealed by the electron microscope in Figure 27·2. This may be generally true of crystalline particles but cannot be true of liquid droplets.

Stabilization of Aerosols

Although the particles of aerosols are inherently unstable, there is some slight evidence that coagulation can be retarded. Droplets surrounding very hygroscopic nuclei are of course more resistant to evaporation, but the problem is as to whether it is possible to diminish the effectiveness of collisions.

The particles of aerosols have sorptive properties and some of these sorbed layers might be effective. Whytlaw-Gray obtained some evidence that certain very thin, liquid films may cause a decreased rate of coagulation. Smirnow and Solntseva [17] asserted that ammonium chloride smoke is stabilized by isoamyl alcohol and that a mist of solar oil is stabilized by water-isoamyl alcohol and by butyric acid. It would be an important advance if stabilized or protected aerosols could be developed. L. Dautrebande states (private communication) that he has achieved this for certain aerosols.

Other properties of aerosols can be altered by sorption; for example, pyrophoric aerosols of iron can be inactivated by surface films of oxide.[18]

Light Scattering by Aerosols

In 1919 R. C. Tolman found that the light scattering from dilute, homogeneous, monodisperse aerosols was proportional to their concentration over a wide range and could be used to measure them.

Effect of Thermal Gradient or of Illumination

It is common knowledge that the cold wall near or behind a heater becomes a depository of dust, whereas there is a dust-free space immediately around hot bodies. This is explained by the unequal bombardment of the particle on the two sides by the Brownian movement of the gaseous medium.

Light has a strong effect upon many aerosols, producing a movement of several centimeters per second, though a few, such as calcium fluoride, calcium sulfate, magnesium oxide, are quite insensitive. The photophoretic effect always occurs in both directions, both away from the light and with some particles traveling toward it. The direction of movement depends upon which side of the particle has the most pronounced temperature gradient, and hence the movement can be quite irregular and may reverse.

Fire-Fighting Fogs

Atomized water issuing at high velocity from special nozzles is frequently used in fire fighting, and its effects are commonly ascribed to cooling of the flames and to stifling of combustion by water vapor. It is the author's experience that these sprays may contain as little as 0.01% of water carried in a blast of air traveling between 40 and 70 miles per hour. The water is therefore so diluted as to have no appreciable effect upon the flame, either by cooling or by suffocating. The water droplets may pass through many feet of flame before they are completely evaporated. Hence the action is merely blowing out the flame like blowing out a candle. It is quite otherwise with glowing wood where the blast has to be directed until enough water has been transported to that surface to quench the glowing charcoal.

Germicidal Aerosols [19]

Airborne infection is a source of cross infections in hospitals and may be a factor in epidemics. Much work has been done in the last few years on its control by nontoxic aerosols. The best known is propylene glycol but even sodium hypochlorite has been tested. One gram of 1% solution atomized into forty million cc of air serves for complete sterilization. Solutions of various detergents, both anion active and particularly cation active, are being studied or employed. Incense smoke and smoke from ignited cardboard soaked with potassium nitrate are both highly effective.

For insect control with DDT oil-aerosols, the optimum size of liquid droplets according to La Mer lies between 2 and 40 microns and may be effective more than a mile downwind under favorable conditions. A brief survey of insecticidal aerosols has recently been given by Nelson and Smith.[20] It has been shown that poultry infection is carried by or on the dust from their pens.

Special Examples of Aerosols

It has been repeatedly suggested that ball lightning consists of a highly charged aerosol.[21]

Dust explosions are caused by reactive dusts such as sugar, soap, starch, coal, or even a petroleum mist.

In December, 1930, a dense fog along the Meuse Valley in Belgium killed 63 people and seriously affected several hundred more. The fog contained sulfur dioxide and sulfuric acid. One fifth of the sulfur came from domestic fireplaces. No matter how efficient a furnace may be in burning all carbonaceous matter and avoiding black smoke, the sulfur oxides remain and can be removed only by washing in an alkaline solution.

Silica dusts are the cause of silicosis. They are profoundly affected [22] by the presence of extraneous substances and by hydrolysis and decomposition of silicates. Chemical changes may occur very rapidly in fine dusts. In general the properties of aerosols can be greatly modified or intensified by suitable admixtures.

The Destruction of Aerosols

Smokes are often merely diluted by allowing the hot gases to escape from a sufficiently tall chimney. At ground level the contamination of the air near the chimney is small but reaches a maximum at a distance of about 10 times the chimney height, but at this distance the smoke is exceedingly dilute.

The most widely used commercial method of destroying aerosols is the Lodge-Cottrell method of electrostatic precipitation.[23] The gases pass over a series of sharp points charged to a high potential. The point discharge electrifies the particles, which are then attracted to and deposited upon an oppositely charged electrode. Valuable residues may thus be recovered. In one cement plant at Riverside, California, the Cottrell precipitator has collected 350,000 tons of cement which would otherwise have been distributed upon the surrounding

citrus groves. W. Witt reported in 1936 that the largest smelter in Montana formerly put into the atmosphere every day one ton each of lead, zinc, copper, and antimony trioxide; about six tons of silica and oxides of iron and aluminum; and over thirteen tons of arsenic. In Los Angeles, California, it has been estimated that various sources, including automobiles, daily put into the air over a thousand tons of sulfurous gases and products.

It may seem that town fogs disperse with abnormal slowness, but the coalescence of the droplets follows the equations already discussed. The number concentration of the droplets is low and the particles are large; hence, their rate of disappearance is slow. If 10^5 particles per cc are present initially, after about $3\frac{1}{2}$ hours there will be only about one half as many, that is 5×10^4 particles per cc. It would take nearly $4\frac{1}{2}$ hours longer to again halve the number of particles to 2.5×10^4 per cc. Meanwhile particles may be added from various sources of pollution.

Other methods of removal are the scrubbing method already mentioned, centrifugal separation, removal by sprays, aggregation of clouds by dusting with dry ice or silver iodide or other nuclei, or coagulation by sound or ultrasonic waves of high intensity. The latter is of theoretical interest because of the several hundredfold increase in size of the particles that almost immediately occurs, followed by coagulation and precipitation. This has been especially studied in a long series of papers by E. Hiedemann. The coagulation is caused by oscillation of the finer particles between node and antinode multiplying collisions. The coagulated particles form double lines in the neighborhood of the nodes.

It is appropriate that this introduction to the literature and main findings of colloid science should have ended with a consideration of aerosols where the main emphasis has necessarily been upon association — the basic and most varied phenomenon of colloid science.

REFERENCES

Aerosols

The standard monograph on this subject is *Smoke: A Study of Aerial Disperse Systems*, by R. WHYTLAW-GRAY and H. S. PATTERSON, Edward Arnold & Co., London (1932) Other collected treatments are: W. E. GIBBS, *Clouds and Smoke*, Blakiston, London (1924); A. WINKEL and G. JANDER, *Schwebstoffe in Gasen, Aersole*, Ahrens Sammlung, F. Enke, Stuttgart (1934); "A General Discussion on Disperse Systems in Gases: Dust, Smoke and Fog," *Trans. Farad. Soc.*, **32**, 1041–1300 (1936)

Text References

1. LANGMUIR, I., Chicago meeting, A.C.S., 1948
2. FINKELSTEIN, L., Chicago meeting, A.C.S., 1948
3. MCBAIN, J. W. and GLASSBROOK, C. I., *J. Am. Chem. Soc.*, **65**, 1908 (1943)
4. LA MER, V. K., SINCLAIR, D. and WILSON, I., Chicago meeting, A.C.S., 1948
5. SVEDBERG, T., *Colloid Chemistry*, A.C.S. Monograph #16, Chemical Catalog Co., Inc., New York (1924)
6. RODEBUSH, W. H., Chicago meeting, A.C.S., 1948
7. NOLAN, J. J., and GUERRINI, V. H., *Trans. Farad. Soc.*, **32**, 1175 (1936)
8. VONNEGUT, B., Chicago meeting, A.C.S., 1948
9. BEISCHER, D. and WINKEL, A., *Z. physik. Chem.*, **176A**, 1 (1936)
10. NIEBURGER, M., Chicago meeting, A.C.S., 1948
11. LANGMUIR, I., Chicago meeting, A.C.S., 1948
12. BROWN, C. E., Bureau of Mines R.I. #3780 (1944)
13. GREEN, H. L., *Phil. Mag.* (VII) **4**, 1046 (1927)
14. GREEN, H. L., *Trans. Farad. Soc.*, **32**, 1091 (1936)
15. GUCKER, F. T., O'KONSKI, C. T., PICKARD, H. B. and PITTS, J. N., *J. Am. Chem. Soc.*, **69**, 2422 (1947)
16. WHYTLAW-GRAY, *loc. cit.*
17. SMIRNOW, L. V. and SOLNTSEVA, V. A., *Coll. J.* (USSR) **4**, 401 (1938)
18. WINKEL, A. and HAUL, R., *Z. Elektrochem.*, **44**, 611 (1938)
19. JENNISON, M. W., in J. ALEXANDER, *Colloid Chemistry*, Vol. V, Reinhold Publishing Corp., New York (1944) p. 1110
20. NELSON, O. A. and SMITH, C. M., in J. ALEXANDER, *Colloid Chemistry*, Vol. VI, Reinhold Publishing Corp., New York (1946) p. 268
21. CAWOOD, W. and PATTERSON, H. S., *Nature*, **128**, 150 (1931)
22. BRISCOE, H. V. A., HOLT, P. F., MATTHEWS, J. W. and SANDERSON, P. M., *Nature*, **139**, 753 (1937)
23. WELCH, H. V., in J. ALEXANDER, *Colloid Chemistry*, Vol. VI, Reinhold Publishing Corp. (1946) p. 274

CHEMICAL AND COMPOUND INDEX